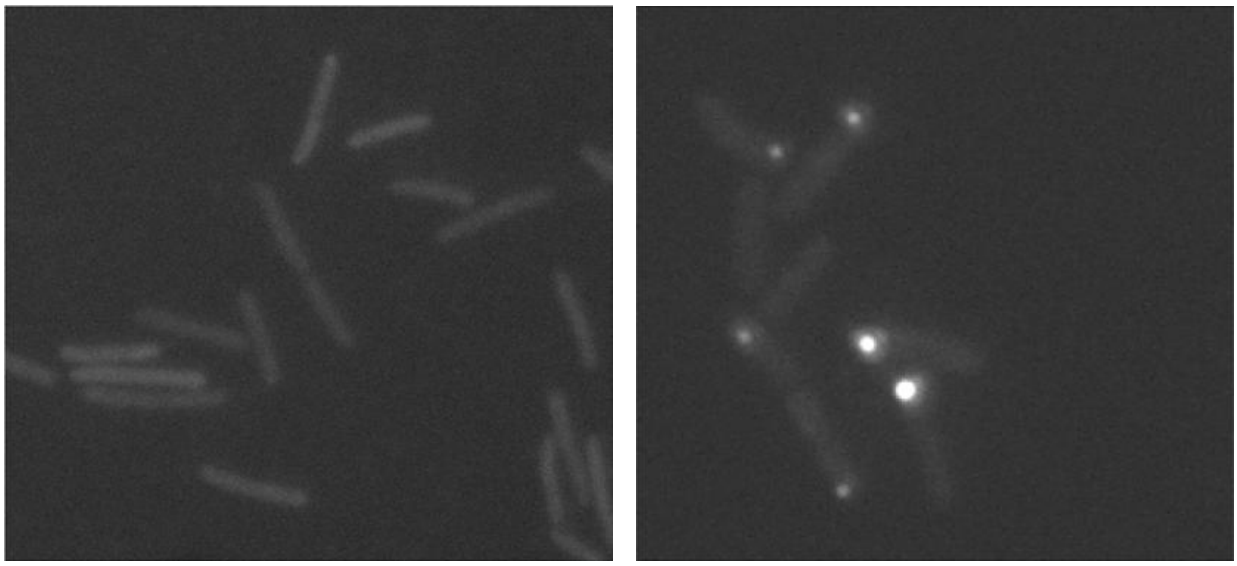
Protein-Tyrosine Phosphorylation in Bacillus Subtilis Signal Transduction

Jers, Carsten

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Protein phosphorylation in *Bacillus subtilis* signal transduction



Carsten Jers

PhD thesis, June 2010

Center for Systems Microbiology
Department of Systems Biology

Protein phosphorylation in *Bacillus subtilis* signal transduction

PhD thesis

by

Carsten Jers

June 2010

Department of Systems Biology
Technical University of Denmark
Kongens Lyngby
Denmark

Frontpage image:

Localisation of enolase in exponential and stationary phase

Preface

This thesis was prepared at Center for Systems Microbiology, DTU Systems Biology at the Technical University of Denmark in partial fulfillment of the requirements for acquiring the PhD degree. It contains the results obtained during my pre-doctoral work from March 2008 to June 2010 carried out mainly at Center for Systems Microbiology but also during a six months stay at AgroParisTech-INRA-CNRS, Thiverval-Grignon, France.

The thesis deals with different aspects of protein phosphorylation in the bacterium *Bacillus subtilis* and is a natural continuation of a ground-breaking phosphoproteome study that significantly increased the number of known phosphorylated proteins. This work has initiated the functional characterisation of these phosphoproteins with a special emphasis on tyrosine phosphorylated proteins.

The study was funded by DTU as a PhD scholarship and was supervised by Professors Ivan Mijakovic and Peter Ruhdal Jensen.

Rødovre, June 14th

Carsten Jers

Summary

Bacteria use protein phosphorylation to regulate all kinds of processes in the cell. In a pioneering gel-free and site specific phosphoproteomics study a number of new serine, threonine and tyrosine phosphorylated proteins in *Bacillus subtilis* were identified and this has provided a platform for further research. Phosphoproteomic studies essentially provide long lists of phosphorylation sites and in this work I have started the next important task: the functional characterisation of the phosphorylation events, focusing on the tyrosine phosphorylated proteins as well as the important transition phase two-component system kinase DegS.

The *B. subtilis* bacterial tyrosine (BY)-kinase PtkA phosphorylates and thereby regulates the activity of two classes of substrates: UDP-glucose dehydrogenases and single-stranded DNA-binding proteins. The phosphoproteomic study identified an additional nine proteins phosphorylated on tyrosine. In this work the nine tyrosine phosphorylated proteins identified in the phosphoproteome study, were tested as substrates of PtkA and the majority of the proteins were phosphorylated by PtkA *in vitro*. In two cases, aspartate semialdehyde dehydrogenase Asd and single-stranded DNA exonuclease YorK, the proteins were activated by PtkA-dependent phosphorylation. Since enzyme activity was not affected in most cases, fluorescent protein tags were used to study the impact of PtkA on localisation of the substrates *in vivo*. Several substrates co-localised with PtkA, and more importantly, the localisation pattern of the proteins enolase, Ldh, YjoA, YnfE, YvyG, Ugd and SsbA was completely altered in Δ*ptkA* background. Interestingly no effects on localisation were observed for YorK and Asd. These results confirm that PtkA-based regulation is achieved by modulating enzyme activity in some cases, but this study also indicates that PtkA in other cases act in terms of ensuring correct cellular localisation of its targets.

The second part of the work was focused on the two-component system DegS/U that is subject to a very complex regulation exerted at both transcription as well as protein level. The cytosolic sensory kinase DegS integrates a wide array of metabolic inputs that modulate its activity. This in turn leads to a finely tuned level of phosphorylation of response regulator DegU that at different phosphorylation levels triggers different sub-regulons. Recently, DegS was found to be

phosphorylated on a serine residue located in its signal sensing domain. In this study it was demonstrated that phosphorylation of DegS could be accomplished by two Hanks type serine/threonine kinases *in vitro*, and this stimulated phosphate transfer to DegU. The consequences of the phosphorylation event were studied *in vivo* using phosphomimetic (S76D) and non-phosphorylatable (S76A) mutants of DegS. In different physiological assays focused on processes regulated by DegU, the phosphomimetic mutant behaved like a strain with intermediate levels of DegU phosphorylation while the non-phosphorylatable mutant behaved like a strain with very low levels of DegU phosphorylation. These results suggest a link between phosphorylation of DegS serine 76 and the level of DegU phosphorylation establishing the phosphorylation event as an input for this two-component system.

In conclusion this thesis work has furthered the understanding of bacterial serine/threonine and tyrosine phosphorylation by revealing a novel mode of action for BY-kinases as well as describing for the first time, a two-component system sensory kinase regulated by serine phosphorylation of its input domain.

Resumé

Bakterier benytter proteinfosforylering til at regulere en masse forskellige processer i cellen. I et banebrydende massespektrometri-baseret fosfoproteomstudie i *Bacillus subtilis* blev et stort antal fosforyleringssteder samt proteiner fosforyleret på serin, threonin og tyrosin identificeret. Dette har skabt et fundament for videre forskning. Resultatet af fosfoproteomstudier er lange lister med fosforyleringssteder, og i dette projekt har jeg påbegyndt det næste vigtige trin: den funktionelle karakterisering af disse fosforyleringer. Specifikt har jeg fokuseret på gruppen af tyrosinfosforylerede proteiner samt den serinfosforylerede sensor kinase i tokomponentsystemet DegS/U, der spiller en vigtig rolle i overgangsfasen mellem eksponentiel og stationær vækstfase.

I *B. subtilis* regulerer BY (Bakteriel tyrosin)-kinasen PtkA to klasser af substrater, UDP-glukose dehydrogenaser samt enkeltstrenget-DNA-bindende proteiner (SSB), ved hjælp af fosforylering. I fosfoproteomstudiet blev der identificeret yderligere ni proteiner fosforyleret på tyrosin. I dette studie blev disse ni proteiner testet som substrater for PtkA, og hovedparten viste sig at blive fosforyleret af PtkA *in vitro*. I to tilfælde, aspartat semialdehyd dehydrogenase Asd og enkeltstrenget-DNA exonuclease YorK, blev proteinerne aktiveret af fosforylering af PtkA. Da enzymaktiviteten i de fleste tilfælde ikke blev påvirket, valgte vi at kigge på betydningen af PtkA i forbindelse med substraternes lokalisering *in vivo*, hvilket blev gjort ved hjælp af proteinsubstrater fusioneret til fluorescerende proteiner. Flere af substraterne co-lokaliserede med PtkA, og vigtigere, lokaliseringsmønstret for proteinerne enolase, YjoA, YnfE, YvyG, Ugd og SsbA var ændret i en Δ*ptkA* stamme. For proteinerne Asd og YorK hvor fosforylering aktiverede enzymaktivitet, så vi ingen ændring i lokalisering. Disse resultater bekræfter at PtkA-baseret fosforylering af dets substrater i nogle tilfælde regulerer enzymaktiviteten, men dette studie indikerer at virkemåden af PtkA i andre tilfælde er at sørge for korrekt lokalisering af dets substrater.

Derudover studerede jeg tokomponentsystemet DegS/U, der er underlagt en meget kompleks regulering på transkriptions- og proteinniveau. Den cytosoliske sensorkinase DegS integrerer et stort antal metaboliske signaler, der leder til en modulering af dets aktivitet. Resultatet af dette er en fintuning af fosforyleringsgraden af DegU, som ved forskellige fosforyleringsgrader aktiverer

transkription af forskellige grupper gener. DegS bliver fosforyleret på en serin i dets signalsensor domæne. I dette studie viste vi, at fosforylering på denne serin kunne udføres af to Hanks-type serin/threonin kinaser *in vitro* og dette stimulerede fosforylering af DegU. Konsekvenserne af fosforylering blev studeret *in vivo* ved hjælp af mutanter der simulerede fosforyleret (S76D) og ufosforyleret (S76A) DegS. I forskellige fysiologiske assays reflekterende processer reguleret af DegU, opførte *degS* S76D mutanten sig som en stamme med moderat niveau af fosforyleret DegU, mens *degS* S76A mutanten opførte sig som en stamme med et lavt niveau af fosforyleret DegU. Disse resultater indikerer en forbindelse mellem fosforylering på serin-76 og fosforyleringsgraden af DegU, og dermed at fosforyleringen af DegS er et aktiverende signal for dette tokomponentsystem.

Dette arbejde har bidraget med en øget forståelse af serin/threonin- og tyrosinfosforylering i bakterier ved at identificere en ny virkemåde for BY-kinaser såvel som ved for første gang at beskrive et tokomponentsystem, der bliver reguleret ved hjælp af fosforylering af dets sensordomæne.

List of publications

1. Jers, C, Pedersen, M.M., Paspaliari, D.K., Schütz, W., Johnsson, C., Soufi, B., Macek, B., Jensen, P.R. and Mijakovic, I. (2010) *Bacillus subtilis* BY-kinase PtkA controls enzyme activity and localisation of its protein substrates. *Mol Microbiol* (Accepted)

2. Jers, C., Søndergaard, E.O., Kobir, A., Jensen, P.R. and Mijakovic, I.(2010) *Bacillus subtilis* two-component system sensory kinase DegS is regulated by serine phosphorylation in its input domain. *Mol Microbiol* (submitted)

Publications not part of this thesis (appendices I-IV)

I. Soufi, B., Jers, C., Hansen, M.E., Petranovic, D. and Mijakovic, I. (2008) Insights gained from bacterial phosphoproteomics. *Biochim Biophys Acta* 1784:186-192

II. Jers, C., Soufi, B., Grangeasse, C., Deutscher, J., and Mijakovic, I. (2008) Phosphoproteomics in bacteria: towards a systemic understanding of bacterial phosphorylation networks. *Expert Rev Proteomics* 5:619-627

III. Miller, M.L., Soufi, B., Jers, C., Blom, N., Macek, B., and Mijakovic, I. (2009) NetPhosBac - a predictor for Ser/Thr phosphorylation sites in bacterial proteins. *Proteomics* 9:116-125

IV. Shi, L., Kobir, A., Jers, C. and Mijakovic, I. (2010) Bacterial protein-tyrosine kinases. *Current proteomics* (In press)

Acknowledgements

This thesis represents the end of my close to four years at the Center for Systems Microbiology as first a master student and now PhD student working on protein phosphorylation in *B. subtilis* in the lab of Ivan Mijakovic.

First of all I would like to thank my supervisors Ivan Mijakovic and Peter Ruhdal Jensen. Special thanks to Ivan for accepting me in his group and for encouraging me to apply for the PhD scholarship. It is difficult to find the words to express my gratitude and respect but I want to thank you for your constantly inspiring guidance, for always finding time to discuss both small and big issues and for your passion and positive nature. I would like to thank Peter Ruhdal Jensen for the continuous support throughout the study and for the many valuable inputs and suggestions especially after Ivan's move to France. Also I would like to thank Flemming G. Hansen for all the valuable help concerning fluorescence microscopy.

Further I would like to express my gratitude towards all the people that I have worked with here at DTU and especially the many people working alongside me in lab 225 for creating a great atmosphere. Too many to name but none forgotten! I would like to thank my old office mates Mette and Jakob and the new ones Anne and Jakob for excellent company. I would like to specifically thank the students Malene, Christina, Dafni, and Mette E. that I had the pleasure of supervising. It was a great joy to see you develop and you did an excellent job that I have also benefitted from. Bo and Elsebeth, thanks for lots of fun days in the lab! Special gratitude goes to Momo. It has been an honour and pleasure getting to know you and further I am indebted to you (and your wife) for feeding me throughout the last four years!

I would also like to thank Ramdane, Amandine and all the other wonderful people at Microbiologie et Génétique Moleculaire in Thiverval-Grignon for making the stay such a memorable one.

Finally I wish to thank Mette, my parents, family and friends for their continuous love and support.

Table of contents

.

Chapter 1

Introduction

INTRODUCTION

The subject of this thesis is the post translational modification protein phosphorylation on serine, threonine and tyrosine residues in *Bacillus subtilis*.

B. subtilis is a Gram-positive, rod-shaped bacterium commonly found in soil. It is widely used in industry for the production of commercially valuable molecules such as enzymes and food additives. Further, *B. subtilis* is a model organism of Gram-positive bacteria, in which extensive studies have contributed to understanding biological phenomena such as sporulation, natural competence and carbon catabolite regulation. With more than 50 years of extensive study it is one of the best understood bacteria with respect to molecular and cell biology.

This work itself is a natural extension of the phosphoproteomic study by Macek and co-workers that identified 78 phosphoproteins and more than doubled the number of known tyrosine phosphorylated proteins in *B. subtilis* (Macek *et al*., 2007). These include proteins that participate in translation initiation, peptide transport, glycolysis, amino acid metabolism, motility, DNA-metabolism, and several proteins with unknown functions (Macek *et al*., 2007). Some of the proteins are encoded by essential *B. subtilis* genes (Kobayashi *et al*., 2003). Other interesting observations are the over-representation of phospho-proteins in carbon metabolism with almost all proteins in glycolysis and several in TCA and pentose phosphate pathway phosphorylated and also the fact that important global transcriptional regulators CodY and DegS are phosphorylated.

This chapter will provide a summary of the known facets of serine/threonine and tyrosine phosphorylation in bacteria in general, followed by an introduction to phosphoproteomics – the study of protein phosphorylation on a global scale. A more detailed description of protein phosphorylation in *B. subtilis* and a summary of the main findings in this work will conclude the chapter.

1.1 Protein phosphorylation in bacteria

Protein phosphorylation is the covalent attachment of a phosphate group to a protein and is a widely employed post translational modification found in all three domains of life, where it plays key roles in regulation and signal transduction.

Historically, protein phosphorylation was first described in *Eukarya* with the demonstration of enzymatic protein phosphorylation in 1954 (Burnett and Kennedy, 1954) and followed up with the work on phosphorylase by Nobel Laureates Edwin Krebs and Edmund Fischer (Fischer and Krebs, 1955). Although there were a few reports of protein phosphorylation in bacteria in the sixties, the general perception emerged that protein phosphorylation on serine, threonine and tyrosine was absent in bacteria (Rubin and Rosen, 1975). The first evidence for the presence of this type of phosphorylation in bacteria came in 1978 with the identification of serine- and threonine-phosphorylated proteins in *Salmonella typhimurium* (Wang and Koshland, 1978) but it was not until the 1990s with the characterisation of the first eukaryotic-like serine/threonine protein kinase, Pkn1, from *Myxococcus xanthus* (Muñoz-Dorado *et al*., 1991) that it was generally accepted that serine/threonine phosphorylation takes place in bacteria. With the growing number of sequenced bacterial genomes encoding eukaryotic-like serine/threonine kinases and phosphatases this is now indisputable (Leonard *et al*., 1998).

The acceptance of tyrosine phosphorylation in bacteria has been longer underway. Although phosphotyrosine was demonstrated as early as 1986 in *Escherichia coli* (Cortay *et al*., 1986) and further established in *Acinetobacter calcoaceticus* in 1990 (Dadssi and Cozzone, 1990) it was not before the purification and characterisation of the first autophosphorylating protein tyrosine kinase Ptk from *Acinetobacter johnsonii* (Grangeasse *et al*., 1997) that protein tyrosine phosphorylation was generally accepted to take place in bacteria. Part of the reason for this is related to the analytical challenges in detection of tyrosine phosphorylated proteins due to low occupancy of phosphorylation sites (Mijakovic *et al*., 2003) and that, unlike serine/threonine phosphorylation, tyrosine phosphorylation is mediated mainly by a bacteria specific kinase showing no resemblance to eukaryotic-like tyrosine kinases (Cozzone *et al*., 2004). The first endogenous substrate of a tyrosine kinase was identified in 2003 and since then the number of characterised kinases and substrates have increased steadily. At present it is clear that serine/threonine and tyrosine phosphorylation affects all parts of bacterial physiology.

In this work the focus is on serine/threonine and tyrosine phosphorylation but bacteria employ at least two other phosphorylation systems; one being the phosphoenolpyruvate-dependent phosphotransferase system (PTS) and the other the two-component systems. Two-component systems involve a sensor kinase and a response regulator. In response to a stimulus, the sensor kinase autophosphorylates on a histidine residue using ATP as phosphate donor. This phosphate

group is subsequently transferred to an aspartate residue on the response regulator. The response regulator is activated by phosphorylation ultimately leading to activation or repression of transcription of specific genes (Stock *et al*., 2000). Although the sensor kinases are often regarded as membrane proteins receiving an external signal, it has been estimated that about 20 % are soluble and sensing a cytoplasmic signal (Mascher *et al*., 2006). The second system is the PTS in which a phosphate group from phosphoenolpyruvate is passed down a chain of specific proteins, all reversibly phosphorylated, and finally transferred to a sugar. The PTS is thus primarily involved in translocation of sugars across the membrane, but phosphorylated forms of some of its components (HPr, EIIA) also have other regulatory roles in the cell (Deutscher *et al*., 2005). Other protein phosphorylation systems however seem to exist, for example the *B. subtilis* protein McsB, formerly thought to be a tyrosine kinase, was recently reported to be a protein arginine kinase (Fuhrmann *et al*., 2009) opening up a new chapter in bacterial protein phosphorylation.

Opposed to the stringent separation of phosphorylation systems there are several examples of cross-talk between phosphorylation systems. The PTS protein HPr is phosphorylated on serine-46 by HPr kinase/phosphorylase. P-Ser-HPr interacts with CcpA and the complex binds to *cre* sites leading to carbon catabolite repression (Deutscher *et al*., 2006) and is linked to virulence in pathogenic strains (Deutscher *et al*., 2005). Recently HPr was shown to be phosphorylated on serine-12 by Hanks type serine/threonine kinase PrkC as well (Pietack *et al*., 2010) Another example of cross-talk comes from recent reports of two-component response regulators phosphorylated by serine/threonine kinases. In *Streptococcus pneumoniae* the kinase StkP phosphorylates the orphan response regulator RitR (Ulijasz *et al*., 2009). The response regulator CovR is phosphorylated on threonine-65 by Stk1 in Group B *Streptococcus* leading to reduced phosphorylation on aspartate-53 and reduced DNA-binding (Lin *et al*., 2009b). A variation of this theme, namely the phosphorylation of a two component sensor kinase, *B. subtilis* DegS, on a serine residue is described in chapter 3.

1.2 Serine/Threonine phosphorylation

Serine/threonine kinases

While serine/threonine phosphorylation in *Eukarya* is mediated mainly by the Hanks type kinases (Hanks *et al*., 1988) a number of different approaches to accomplish this task has evolved in bacteria. At least four different classes of protein serine/threonine kinases exist of which two

contain both modifying and demodifying activities in a single polypeptide. The first is isocitrate dehydrogenase kinase/phosphatase (LaPorte and Chung, 1985) that phosphorylates isocitrate dehydrogenase on a serine residue. Phosphorylation inactivates the enzyme and thus regulates carbon flux (Borthwick *et al*., 1984). The second is HPr kinase/phosphorylase that via phosphorylation of HPr (and Crh) regulates carbon catabolite repression. This protein employs the Walker A motif for nucleotide binding and its activities are regulated by intracellular metabolites. An interesting feature of this protein is that the demodification is done using inorganic phosphate instead of H_2O for nucleophilic attack and hence it is a phosphorylase opposed to the more normal phosphatase activity (Mijakovic *et al*., 2002). Another class of kinases found in many, especially, Gram-positive bacteria exhibits a weak sequence homology to two-component system histidine kinases. These kinases have been well studied in *B. subtilis* represented by SpoIIAB (Min *et al*., 1993). Although undetectable at the sequence level the crystal structure demonstrated some resemblance with Hanks type kinases in the active site (Masuda *et al*., 2004). Finally a number of kinases bearing resemblance to eukaryotic-like (Hanks type) kinases have been identified. While the other classes of protein kinases seem to play very specific roles with only one or few substrates the Hanks type kinases are thought to play more diverse roles (Pietack *et al*., 2010). An interesting observation is that the catalytic domain in many instances is fused to other domains thought to reflect the regulatory role of the protein (Krupa and Srinivasan, 2005). These domains include enzymatic, protein-protein interaction and signal sensing domains. Many of the kinases are membrane proteins and it has been suggested that they could function as receptor like kinases that bind extracellular signals. Such a role was recently shown for *B. subtilis* Hanks type kinase PrkC that contains PASTA (penicillin-binding protein and serine/threonine kinase associated) domains. In spores, the binding of muropeptides (released from the cell wall of vegetative cells) by the PrkC PASTA domain represents an activating signal that leads to spore germination (Shah *et al*., 2008).

Serine/threonine phosphatases

Aside from the phosphatase and phosphorylase activities of isocitrate dehydrogenase kinase/phosphatase and HPr kinase/phosphorylase respectively, bacterial serine/threonine phosphorylation is counteracted by two classes of eukaryotic-like phosphatases: the phosphoprotein phosphatase (PPP) and the protein phosphatase Mg^{2+} or Mn^{2+} dependent (PPM) families. These families exhibit rather similar architecture of catalytic domains but are unrelated at the sequence

level (Barford, 1996). While the catalytic domains are similar, different regulatory domains, as mentioned above for kinases, are found in the individual phosphatases reflecting their regulatory roles. PPM phosphatase RsbP in *B. subtilis* for example contains an N-terminal PAS (Per-ARNT-Sim) domain involved in sensing energy stress (Vijay *et al*., 2000)

Physiological role and protein substrates

The regulatory function of serine/threonine phosphorylation is very diverse and has the potential to regulate practically all kinds of processes in the cell. Phosphoproteomic studies showed an over-representation of phospho-proteins in carbon metabolism and especially among glycolytic enzymes (Soufi *et al*., 2008a). At this point however the role(s) of these phosphorylation events are largely unknown but the very first identified bacterial protein was isocitrate dehydrogenase found in the tricarboxylic acid (TCA) cycle that is phosphorylated on a conserved serine residue and thereby inactivated. The phosphorylated serine residue is located in the active site and the negative charge abolishes binding of the substrate isocitrate by preventing hydrogen bonding and creating electrostatic repulsion (Dean *et al*., 1989; Dean and Koshland, 1990). Inactivation of isocitrate dehydrogenase enables the bacterium to bypass the TCA cycle via the glyoxylate shunt when grown on acetate and thus regulates carbon flux. Further, carbon catabolite repression is regulated by serine phosphorylation of HPr. P-Ser-HPr interacts with transcriptional regulator CcpA and this complex binds catabolite response elements (*cre* sites) thereby regulating a large number of genes involved in sugar metabolism and utilisation (Deutscher *et al*., 2006). P-Ser-HPr has also been linked to pathogenicity by virtue of controlling transcription of virulence related genes (Deutscher *et al*., 2005).

In virulence, kinases and phosphatases are heavily used as secreted effector proteins that during different steps of infection scramble host cell signalling pathways (Jers *et al*., 2008). Phosphorylation can however also affect the secretion process itself. In *Pseudomonas aeruginosa* threonine phosphorylation of serine/threonine kinase PpkA mediates protein-protein interaction with a FHA (Forkhead-associated)-domain and thereby controls the activity of a type VI secretion system (Mougous *et al*., 2007). The FHA domain binds threonine phosphorylated peptides and is known to mediate phosphorylation dependent protein-protein interactions. FHA domains have been predicted in many but not all bacterial genomes and in many cases their cellular role remains

elusive (Pallen *et al*., 2002). A variation of this theme was recently reported in regulation of the TCA cycle in *Corynebacterium glutamicum*. The FHA domain protein Odh1 binds and inhibits 2-oxoglutarate dehydrogenase. Odh1 is phosphorylated on an N-terminal threonine residue and this leads to a major conformational change where the N-terminal phospho-threonine binds to the C-terminal FHA domain thereby abolishing binding to 2-oxoglutarate dehydrogenase and relieving inhibition (Barthe *et al*., 2009).

In many Gram-positive bacteria activity of the alternative sigma factor B is controlled by a partner switching cascade where phosphorylation controls protein-protein interaction specificity. In unstressed cells the sigma factor is sequestered by an anti-sigma factor (a serine kinase) while the anti-anti sigma factor is phosphorylated rendering it unable to bind the anti-sigma factor. In response to stress the anti-anti-sigma factor is dephosphorylated, it then binds the anti-sigma factor while the sigma factor is released and activates gene expression. Identified roles of this system includes the stress response in *B. subtilis* (Yang *et al*., 1996), regulation of biofilm formation in *Staphylococcus epidermidis* (Knobloch *et al*., 2004), type III secretion in *Bordetella* species (Mattoo *et al*., 2004) and the developmental process sporulation in *B. subtilis* (Alper *et al*., 1994).

M. xanthus is a model organism for studying development. Upon starvation, some 100.000 cells move towards an aggregation center and form a multicellular fruiting body. About 10 % of the cells will differentiate into metabolically dormant myxospores while another 80 % lyse thereby providing nutrients to spores and fruiting body forming cells. *M. xanthus* encodes more than 100 Hanks type kinases of which several have been characterised and shown to play a regulatory role in fruiting body formation (Inouye and Nariya, 2008). Interestingly, a kinase cascade has been identified in this organism where serine/threonine kinase Pkn8 phosphorylates Pkn14 that in turn phosphorylates an important transcription factor MrpC (Nariya and Inouye, 2005). In the last decade "undomesticated" (non-lab) strains of *B. subtilis* have been shown to exhibit social behaviour and development similar to *M. xanthus* (Branda *et al*., 2001) and this work indicates that serine phosphorylation could play a (so far largely undefined) role in this process (see chapter 3).

While not an exhaustive survey of the known roles of serine/threonine phosphorylation in bacteria, the presented examples demonstrate the versatility with which protein phosphorylation dependent regulation is exerted. This ranges from modulating enzyme activity, mediating protein-protein interaction and regulating gene expression thereby playing important roles in all parts of bacterial physiology.

1.3 Tyrosine phosphorylation

While serine/threonine phosphorylation is determined by several different classes of proteins, tyrosine phosphorylation appears to be conducted primarily by one class: the BY (bacterial tyrosine) kinases. The exceptions to the rule being the kinase DivL from *Caulobacter crescentus* that shows homology to two-component system sensory kinases but autophosphorylates on a tyrosine residue located in place of the conserved histidine residue. Upon autophosphorylation on tyrosine, the response regulator CtrA is phosphorylated on aspartate probably by phosphotransfer thereby regulating cell division and differentiation (Wu *et al*., 1999). Two proteins resembling eukaryotic-like tyrosine kinases MasK from *M. xanthus* and *P. aeruginosa* WaaP have been identified. In both cases the (putative) role as protein kinase remains unclear but MasK interacts with MglA and thereby apparently regulate motility and development (Thomasson *et al*., 2002) while WaaP is an essential sugar kinase involved in lipopolysaccaride synthesis (Zhao and Lam, 2002). Recently a novel tyrosine kinase belonging to the HAD (haloacid dehydrogenase-like hydrolase) superfamily was identified in *Mycobacterium tuberculosis*. The kinase autophosphorylates and phosphorylates the secreted virulence factor tyrosine phosphatase PtpA (Bach *et al*., 2009).

The BY-kinase

Sequence analysis early made it clear that BY-kinases represent a novel class of tyrosine kinases not found in eukaryotes. They use the Walker motifs A, A' and B originally found in nucleotide-binding proteins for active site, thus displaying higher resemblance to ATPases such as MinD and Soj (Mijakovic *et al*., 2005a), and further contain a tyrosine-rich cluster in the C-terminus. BY-kinases possess a transmembrane domain and an intracellular catalytic domain. In *Proteobacteria*, these domains are situated in a single protein, while they are found as two separate proteins in *Firmicutes*, encoded by the same operon (Cozzone *et al*., 2004) (figure 1.1). In the latter case, the transmembrane modulator and the cytosolic kinase communicate through a specific helix-helix interaction that modulates kinase activity (Soulat *et al*., 2006). The transmembrane modulator is considered to function as a sensor domain that upon sensing a presently unknown signal can trigger kinase activity (Grangeasse *et al*., 2007).

A feature of the BY-kinases is that they autophosphorylate at tyrosine residues located in the C-terminus. The position of tyrosines in the tyrosine-rich cluster is not conserved in different bacteria, and studies in *E. coli* indicate that it is the overall level of phosphorylation rather than a specific combination that dictates the biological role (Paiment *et al*., 2002). BY-kinases of *Proteobacteria* furthermore autophosphorylate on a tyrosine positioned between Walker motifs A' and B as part of a two-step activation process (Grangeasse *et al*., 2002). In *B. subtilis* it has been shown that autophosphorylation does not affect the kinase activity *in vitro* (Mijakovic *et al*., 2003), but structural studies have recently demonstrated an important role of autophosphorylation *in vivo*.

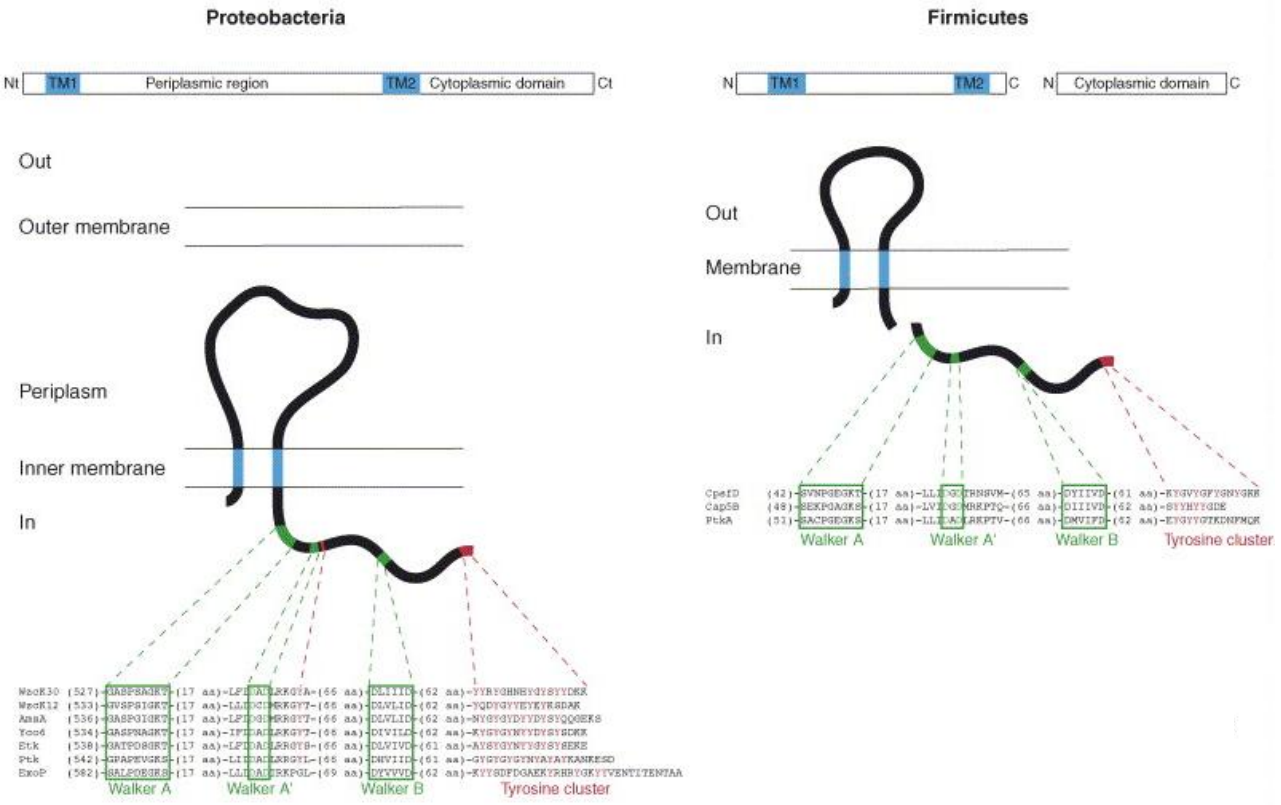


Figure 1.1: BY-kinases in *Proteobacteria* and *Firmicutes*. The BY-kinase consists of a transmembrane modulator domain and a kinase domain. In *Proteobacteria* the domains are found in a single polypeptide while in *Firmicutes* they are split in two polypeptides. The catalytic domain constituted by the Walker A, A' and B motifs (green) as well as transmembrane segments (blue) and tyrosine rich cluster subject to autophosphorylation (red) is presented (Figure adapted and modified from Grangeasse *et al*., 2007).

The first structural insights came from the low resolution structure (14 Å) of *E. coli* BY-kinase Wzc in complex with the capsular polysaccharide translocase Wza revealing a Wzc tetramer that oligomerises via the periplasmic parts with non-interacting kinase domains protruding into the

cytosol. The structure indicated that Wzc regulates export by triggering an open/active conformation of Wza (Collins *et al*., 2007).

The recent publication of two BY-kinase structures, Etk from *E. coli* (Lee *et al*., 2008) and CapA/B from *Staphylococcus aureus* (Olivares-Illana *et al*., 2008) provided insights into the functional role of several conserved features. In a biochemical study it was demonstrated that Etk homologue Wzc is activated through a two-step mechanism involving a conserved tyrosine in the active site and a C-terminal tyrosine cluster (Grangeasse *et al*., 2002). The structure data combined with biochemical and *in silico* analysis revealed a novel mechanism by which the first activation steps proceeds. The conserved tyrosine side chain points into the active site thereby blocking activity. Upon autophosphorylation the negatively charged phosphotyrosine rotates out of the active site and is stabilised by interaction with an arginine residue (Lee *et al*., 2008; Lu *et al*., 2009). In *Firmicutes* the BY-kinase is separated in a cytosolic kinase and a transmembrane modulator that activates the kinase. This activation proceeds, at least in part, by the fact that the modulator protein helps to stabilise ATP binding by hydrophobic sandwich interaction between the adenine ring and a phenylalanine residue of the modulator protein. For CapB, the unphosphorylated kinase monomers, as opposed to Wzc, associate in an octameric ring structure located to the membrane via interaction with its transmembrane modulator CapA (figure 1.2). Notably, a tyrosine residue in the tyrosine cluster is bound in the active site of the neighbouring subunit suggesting an intermolecular phosphorylation mechanism. The fact that no tyrosine residue in the cluster is preferentially phosphorylated reflects a high flexibility in that region. An important conclusion from the study was that autophosphorylation of CapB is expected to induce dissociation of the ring structure, while maintaining the interaction between individual CapB-CapA couples, allowing kinase monomers to phosphorylate its endogenous substrates (Olivares-Illana *et al*., 2008). It has previously been suggested that the different architecture in *Firmicutes* could reflect that these kinases might dissociate from the modulator and phosphorylate its endogenous substrates possibly under control of alternative modulator proteins (Mijakovic *et al*., 2005a) and this is substantiated by data from this work. Localisation data indicated that the kinase monomers upon dissociation of the octameric ring dissociates from the transmembrane modulator which would free them to phosphorylate or possibly interact with its substrates (see chapter 2). The structures also provided a rationale for the substrate specificity of tyrosine vs. serine/threonine kinases and the difference between BY-kinases and the structurally highly similar ATPases such as MinD (Lee and Jia, 2009).

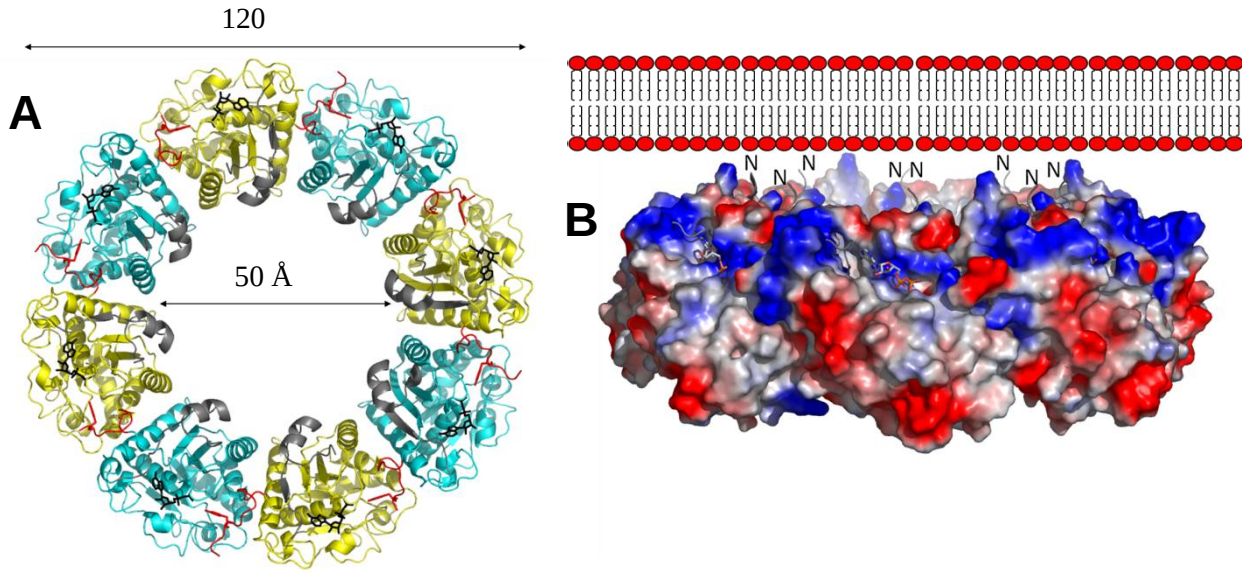


Figure 1.2: Structure of an unphosphorylated *Firmicute* BY-kinase octamer. The structure of BY-kinase CapB from *S. aureus* demonstrated a ring-shaped octamer that is localised to the membrane via interaction with its transmembrane modulator. (A) The ring-shaped octameric structure with individual subunits coloured in yellow and blue. A tyrosine from the tyrosine rich cluster points into the active site of the neighbouring subunit (red). (B) Orientation of the octamer in relation to cell membrane. (Figure adapted and modified from Olivares-Illana *et al*., 2007).

Protein tyrosine phosphatases

In order for tyrosine phosphorylation to be a regulatory device, protein tyrosine phosphatases antagonizing the kinase activity are needed. In bacteria, three different classes of tyrosine phosphatases have been found: two families similar to eukaryal class I (classic) and class II low-molecular-weight-protein tyrosine phosphatases (LMW-PTPs) which both contain a conserved cysteine residue in the catalytic site. The third class belongs to the polymerase and histidinol phosphatase (PHP) family (Morona *et al*., 2002), and are manganese-dependent phosphotyrosine phosphatases so far only described in *Firmicutes* (Grangeasse *et al*., 2007).

The classic phosphatases primarily function as virulence factors of pathogens that upon secretion dephosphorylate host proteins, thereby interfering with host signal transduction networks. They are therefore not of importance with respect to regulation of bacterial cellular functions (Grangeasse *et al*., 2007). The LMW-PTPs seem to play different roles in *Firmicutes* and *Proteobacteria* respectively and this is reflected in the genomic organisation of genes encoding BY-kinases and phosphatases. In *Proteobacteria* the genes encoding LMW-PTPs are found immediately upstream

of the BY-kinases and they have been shown to efficiently dephosphorylate these thereby regulating their biological function. In *Firmicutes* however, the LMW-PTP encoding genes are located elsewhere in the genome. Their role in *Firmicutes* is presently not clear although studies on *B. subtilis* LMW-PTPs indicate that they play a role in stress resistance (Musumeci *et al*., 2005). One of these, YwlE, was previously identified as the cognate phosphatase of McsB, a kinase previously believed to be a tyrosine kinase but recently categorised as an arginine kinase (Kirstein *et al*., 2005; Fuhrmann *et al*., 2009). The PHP type phosphatase found only in *Firmicutes* is encoded immediately up- or downstream of the BY-kinase and modulator genes. They dephosphorylate the BY-kinases and their substrates and therefore play a role similar to LMW-PTPs in Gram-negative bacteria (Grangeasse *et al*., 2007).

Physiological role and protein substrates

The first physiological role attributed to BY-kinases was related to the production of exopolysaccharide where it functions as a co-polymerase. This process has been extensively studied in *E. coli* that encodes two BY-kinases Wzc and Etk. Wzc in involved in the synthesis of the exopolysaccharide colanic acid (Vincent *et al*., 2000) while Etk, present in some *E. coli* strains, is required for the formation of the group 4 capsule (Peleg *et al*., 2005). Synthesis of colanic acid as well as correct size of the polymer requires both the phosphorylated and the unphosphorylated form of Wzc (Obadia *et al*., 2007). Wzc also phosphorylates and thereby activates the enzyme Ugd that is involved in production of precursors for colanic acid (Grangeasse *et al*., 2003). BY-kinases have been implicated in exopolysaccharide and capsule synthesis in a number of other bacteria. In *S. pneumoniae* CpsD is required for correct capsule synthesis and it is essential for virulence (Morona *et al*., 2003; Morona *et al*., 2004). In *Burkholderia cepacia* deletion of tyrosine phosphatase BceF abolishes exopolysaccharide production and deletion of either BceF or BY-kinase BceD inhibits biofilm formation (Ferreira *et al*., 2007) and several other examples exist (Grangeasse *et al*., 2007).

In 2003 the first endogenous substrates of BY-kinases were identified. Tyrosine phosphorylation was shown to affect heat shock response in *E. coli* where the principal heat shock sigma factor RpoH is inactivated by phosphorylaon of BY-kinase Etk (Klein *et al*., 2003). In the same study the anti-sigma factor RseA that binds and thereby sequesters sigma factor RpoE was found to be phosphorylated on tyrosine presumably altering its binding affinity towards RpoE. In parallel the enzyme Ugd (UDP-glucose dehydrogenase) that participates in production of exopolysaccaride was

shown to be phosphorylated by PtkA in *B. subtilis* and Wzc in *E. coli* thereby stimulating its enzyme activity (Mijakovic *et al*., 2003; Grangeasse *et al*., 2003). *E. coli* Ugd represents the only example of a protein modified by two different BY-kinases, Etk and Wzc, affecting its physiological role in antibiotic resistance and production of exopolysaccharide (Lacour *et al*., 2008). Analysis of the activation mechanism of *B. subtilis* Ugd suggested that an active site conformational change is mediated by interaction between phosphotyrosine and a lysine residue but being that the lysine residue is not conserved in other Ugds suggests that the mode of activation might differ in other organisms (Petranovic *et al*., 2009). Although the individual phosphorylation events do not seem to be strongly conserved, the general theme of phosphorylation-based activation of enzymes in sugar polymer synthesis is widespread. Identified events include phosphoglycosyl transferase EpsE phosphorylated by BY-kinase EpsD in *Streptococcus thermophilus* (Minic *et al*., 2007), UDP-acetyl-mannosamine dehydrogenase Cap5O phosphorylated by BY-kinase Cap5B2 (Soulat *et al*., 2007) and undecaprenylphosphate glycosyltransferase (WcaJ) phosphorylated on tyrosine in *Klebsiella pneumoniae* (Lin *et al*., 2009a). Next, phosphorylation of the single-stranded DNA binding proteins were discovered in *B. subtilis*, *E. coli* and *Streptomyces coelicolor* and in *B. subtilis* phosphorylation leads to an increased affinity for single-stranded DNA (Mijakovic *et al*., 2006).

With the recent developments in mass spectrometry-based, gel-free bacterial phosphoproteomics the number of identified proteins phosphorylated on tyrosine has increased dramatically linking tyrosine phosphorylation to all parts of bacterial physiology. Next important steps involve the characterisation of these phosphorylation events.

1.4 Phosphoproteomics – the systematic approach

As early as 1986 it was demonstrated that *E. coli* contains more than 100 phosphoproteins by two-dimensional (2D) gel electrophoresis of radiolabelled proteins (Cortay *et al*., 1986) and later, the 2D-gel approach, has been used to study protein phosphorylation on a global scale identifying 28 and 29 phosphoproteins in *B. subtilis* (Eymann *et al*., 2007; Lévine *et al*., 2006) and 41 phosphoproteins in *C. glutamicum* (Bendt *et al*., 2003). While no phosphorylation sites were identified in these studies a combination of phosphoprotein enrichment and 2D gel electrophoresis followed by mass spectrometry analysis of trypsin digested spots enabled identification of 58

phosphopeptides from 36 proteins including three on tyrosine in *Campylobacter jejuni* (Voisin *et al*., 2007). A characteristic of these studies have been the analysis of phosphorylation patterns in various conditions effectively demonstrating that protein phosphorylation is a highly dynamic process. Recently this approach allowed detection of 63 phosphoproteins and identification of 16 sites from *Mycoplasma pneumoniae* (Schmidl *et al*., 2010).

While the 2D gel approach has certainly moved the field of bacterial protein phosphorylation there are some technical problems with the approach. A problem general to 2D gel electrophoresis is the low recovery of hydrophobic proteins such as membrane proteins, but this can at least partially be overcome by special solubilisation and analysis protocols (Eymann *et al*., 2004). Another problem is the coverage that, by employing several procedures, is about half of the proteins produced (Wolff *et al*., 2006). Further, the wide range of occupation of phospho-sites (from less than 1 % to more than 90 %) complicates the detection. Occupation of tyrosine residues is often very low and only very few proteins phosphorylated on tyrosine have been reported (Mijakovic *et al*., 2003).

Phosphorylation research has been hampered by the difficulties of identifying phosphorylated proteins and more pronounced identification of phosphorylation sites, but in the last couple of years a number of mass spectrometry based phosphoproteomic studies have been conducted (Macek *et al*., 2007; Macek *et al*., 2008; Soufi *et al*., 2008b; Ravichandran *et al*., 2009; Lin *et al*., 2009a). This approach, in short, involves digestion of crude extracts with endoprotease (e.g. trypsin) followed by phosphopeptide enrichment. The mixture is then separated by liquid chromatography that is coupled to a high resolution mass spectrometer. These studies allowed identification of phosphoproteins and in most cases the phosphorylation site could be determined. One of the advantages of 2D gel based phosphoproteomics is the relative ease with which changes in e.g. different growth conditions can be analysed, but such studies are also feasible with MS. One approach is stable isotope labelling by amino acids in cell culture (SILAC) that enables comparison of two strains (e.g. wild type and kinase or phosphatase mutant) or one strain cultured in two different growth conditions (Ong *et al*., 2002).

These pioneering studies have dramatically increased the number of known phosphoproteins but the fact that previously identified phosphoproteins in *B. subtilis* were not detected led to the hypothesis that the identified proteins might only represent the tip of the iceberg. In most studies, phosphopeptides were enriched by strong cation exchange and TiO_2 chromatography. Lin and co-workers in parallel with this approach also used phospho-tyrosine specific antibodies for enrichment

leading to identification of an additional 17 tyrosine phosphorylated peptides out of 24 in total (Lin *et al*., 2009a), indicating that a substantial part of the phosphoproteome pass undetected. Building on this notion, a systematic mapping of phosphorylation sites among ribosomal proteins in *E. coli* identified 24 phosphoproteins of which several were phosphorylated on multiple residues while only five of these were identified in the phosphoproteome (Soung *et al*., 2009; Macek *et al*., 2008).

The recent site-specific phosphoproteomes have greatly extended the list of known phosphoproteins and sites. It has become apparent that an over-representation of phosphoproteins is found among house-keeping pathways and central carbon metabolism but so far no role has been attributed for this. It is becoming increasingly clear that there is a very low conservation of phosphorylation sites and it is thought that this might reflect the adaptation to different ecological niches (Soufi *et al*., 2008b).

These phosphoproteome studies generate long lists of phosphoproteins, and next important steps will involve identifying kinases and phosphatases acting on the proteins and the physiological roles of the phosphorylation events. The first point should in theory be possible to do by comparing phosphoproteomes of wild type and kinase/phosphate knockout strains. Interestingly, a recent study applied this approach in *M. pneumoniae* with the surprising conclusion that only five of 63 phosphoproteins are targeted by the two known kinases pointing towards the presence of novel classes of kinases (Schmidl *et al*., 2010). In *B. subtilis* the targets of serine/threonine kinase PrkC and phosphatase PrpC are currently being identified and this will indicate the generality of this conclusion. Secondly, the dynamics of phosphorylation should be studied, that is to say, how the phosphoproteome changes in response to e.g. growth conditions and growth phase. This can be done on a global scale using approaches such as SILAC and could in some cases probably indicate the physiological roles of phosphorylation events. Thirdly, a daunting task will be to identify the physiological roles of the phosphorylation events, a task that this thesis work has been devoted to.

An alternative to experimental identification of phosphorylation sites is *in silico* prediction. Evaluated on the published phosphoproteome of *B. subtilis* the predictor NetPhos trained on eukaryotic data demonstrated a low performance on bacterial proteins (Soufi *et al*., 2008a). This led to the development of a bacteria-specific version NetPhosBac that outcompeted the predictors dedicated eukaryotic proteins as well as DISPHOS (bacteria) and was validated experimentally (Miller *et al*., 2009). NetPhosBac is currently dedicated serine and threonine phosphorylation due to

the limited number of known tyrosine phosphorylation sites, but it is only a question of time before enough data for a tyrosine specific version is available.

1.5 *Bacillus subtilis* as a model organism

B. subtilis is a Gram-positive model organism in which studies have contributed to understanding biological phenomena such as sporulation, natural competence and carbon catabolite regulation. Of importance in the context of this work it is also one of the most studied prokaryotes with respect to protein serine/threonine/tyrosine phosphorylation and several phosphorylation "modules", especially partner switching in sporulation and stress response, are well understood. *B. subtilis* further contains nine characterised or putative kinases and nine phosphatases. With the phosphoproteome studies in *B. subtilis* more than 100 phosphoproteins have been identified and in most cases the physiological roles of these phosphorylation events remain unknown.

My work on the serine phosphorylation of two-component system sensory kinase DegS led me to also study social behaviour in *B. subtilis* and a paragraph in this introduction is dedicated to this fascinating topic.

Serine/threonine/tyrosine kinases and phosphatases in *B. subtilis*

Protein phosphorylation is controlled by kinases and phosphatases and a number of characterised as well as putative protein kinases and phosphatases belonging to distinct classes have been identified in *B. subtilis* (table 1.1).

These include the well characterised kinases SpoIIAB, RsbT and RsbW that belong to the histidine kinase like family but phosphorylate its substrates on serine. These kinases are involved in partner switching and control transcription in sporulation and stress response respectively and are antagonised by the transmembrane phosphatase SpoIIE and the cytosolic phosphatases RsbX, RsbU and RsbP that all belong to the PPM class. RsbP additionally contains an N-terminal PAS domain, a signal sensing domain.

The class of Hanks type kinases are represented by the transmembrane kinases PrkC and YabT and the cytosolic kinase YbdM (PrkD). PrkC harbours the PASTA domain that is involved in activation

Table 1.1: Serine, threonine and tyrosine kinases and phosphatases in *B. subtilis*. Characterised and putative kinases and phosphatases with information on specificity, substrates and regulated processes.

Name	Specificity	Substrates*	Regulated processes
Kinases			
PrkC	Ser/Thr	EF-G (FusA)[1], YezB[2], CpgA[2], EF-TU[2], YwjH[3], GlnA[3], Icd[3], AlsD[3], HPr[3]	Sporulation, biofilm and germination
YbdM (PrkD)	Ser/Thr	---	Putative kinase /unknown function
YabT	Ser/Thr	---	Putative kinase / unknown function
PrkA	Ser	60 kDa protein[4]	Putative kinase
SpoIIAB	Ser	SpoIIAA[5]	Sporulation (control of sigma F activity)
RsbT	Ser	RsbRA[6], RsbS[6]	Stress (control of sigma B activity)
RsbW	Ser	RsbV[7]	Stress (control of sigma B activity)
PtkA	Tyr	Ugd[8], TuaD[8], SsbA[9], SsbB[9]	Exopolysaccharide synthesis?, DNA metabolism
EpsB	Tyr	---	Putative kinase / Exopolysaccharide synthesis
Phosphatases			
PrpC	Ser/Thr	YezB[2], CpgA[2], EF-TU[2], PrkC[10], EF-G (FusA)[10], HPr[11]	Sporulation, biofilm and germination
SpoIIE	Ser	SpoIIAA[12]	Sporulation (control of sigma F activity)
RsbX	Ser	RsbS[13], RsbRA[14]	Stress (control of sigma B activity)
RsbU	Ser	RsbV[13]	Stress (control of sigma B activity)
RsbP	Ser	RsbV[15]	Stress (control of sigma B activity)
PtpZ	Tyr	PtkA[8], Ugd[8], TuaD[8], SsbA[9], SsbB[9]	Exopolysaccharide synthesis?, DNA metabolism
YwlE	Tyr/Arg?	McsB[16], McsA[16], CtsR[16]	Stress
YfkJ	Tyr	---	Survival of ethanol stress
PrpE	Tyr	---	Spore germination
Kinase/phosphorylase			
HPr K/P	Ser	HPr[17], Crh[17]	Carbon catabolite repression

***Note that some substrates are inferred solely from *in vitro* phosphorylation assays. Phosphorylation events have been reported in the following: [1] Shah and Dworkin, 2010; [2] Absalon *et al*., 2009; [3] Pietack *et al*., 2010; [4] Fischer *et al*., 1996; [5] Min *et al*., 1993; [6] Kim *et al*., 2004; [7] Dufour and Haldenwang, 1994; [8] Mijakovic *et al*., 2003; [9] Mijakovic *et al*., 2006; [10] Gaidenko *et al*., 2002; [11] Singh *et al*., 2007; [12] Arigoni *et al*., 1996; [13] Yang *et al*., 1996; [14] Chen *et al*., 2004; [15] Brody *et al*., 2001; [16] Kirstein *et al*., 2005; [17] Martin-Verstraete *et al*., 1999.**

of the kinase via binding of muropeptides (Shah *et al*., 2008). PrkC phosphorylation is countered by the cytosolic PPM phosphatase PrpC and the corresponding genes are encoded in the same operon (Iwanicki *et al*., 2005).

Two BY-kinases, PtkA and EpsB, exist in *B. subtilis*. PtkA has been well characterised and several substrates have been identified (Mijakovic *et al*., 2003; Mijakovic *et al*., 2006). When expressed in

E. coli EpsB becomes phosphorylated (either autophosphorylation or phosphorylation by *E. coli* kinase) but so far its role as kinase remains elusive and no substrates have been identified (Mijakovic *et al*., 2003). PtkA phosphorylation requires interaction with a modulator protein TkmA and is antagonised by PHP class tyrosine phosphatase PtpZ. Two LMW-PTPs YwlE and YfkJ implicated in stress resistance have been identified (Musumeci *et al*., 2005). One of these, YwlE, however dephosphorylates the arginine kinase McsB, raising question as to whether they are truly tyrosine phosphatases (Kirstein *et al*., 2005; Fuhrmann *et al*., 2009).

Another kinase employing the Walker motifs is HPr kinase/phosphorylase that targets HPr and regulates carbon catabolite control (Deutscher *et al*., 2006). The Walker motif containing protein PrkA does not autophosphorylate but was reported to phosphorylate a 60 kDa protein on serine (Fischer *et al*.,1996) and a *prkA* mutant exhibits defects in spore formation (Eichenberger *et al*., 2003). More evidence would be needed to confirm PrkA to be a protein kinase.

Finally, a PPP type phosphatase PrpE exists. This class usually dephosphorylates phospho-serine and –threonine, but was shown to target tyrosine phosphorylated peptides (Iwanicki *et al*., 2002). Its activity is cell cycle dependent (Sandler and Keynan, 2007) and a role in spore formation has been proposed (Hinc *et al*., 2006).

Two putative kinases, YxaL and YrzF, mentioned in the literature were omitted from table 1.1. YrzF appears to be a fragment of a Hanks type kinase showing homology to conserved regions I, II and III (Leonard *et al*., 1998). The conserved domain database identifies similarity to the Aminoglycoside 3'-phosphotransferase (APH) and Choline Kinase (ChoK) family and hence it is probably unlikely to be a protein kinase (Marchler-Bauer *et al*., 2009)

YxaL contains several Pyrroloquinoline quinone (PQQ) domains but does not show homology to known kinase domains. It has nevertheless been associated with such a function (Pietack *et al*., 2010) and in fact *E. coli* homologue YfgK was shown, using phospho-serine/threonine specific antibodies, to autophosphorylate in the presence of PQQ. It is further a major component of the secretome (Voigt *et al*., 2009). Overall, it seems unlikely that it should be a protein kinase, but to rule it out, it should be tested, at least *in vitro*, whether YxaL, as YfgK, can autophosphorylate in presence of PQQ.

In most cases the responsible kinase(s) and phosphatase(s) targeting the about 100 known phosphoproteins in *B. subtilis* remain unknown. It is generally believed that the HPr kinase and the

kinases SpoIIAB, RsbT and RsbW involved in controlling activity of sigma factors play defined roles and as such are unlikely to phosphorylate the many newly identified phopshoproteins (Pietack *et al*., 2010). With the relative shortage of kinases this would indicate that the remaining kinases would phosphorylate on average 15-20 proteins or that many so far unknown kinases exist. At present there are evidence indicating the validity of both hypotheses.

Serine/threonine phosphorylation in *B. subtilis*

The phosphorylation based partner switching especially in stress response is the best understood serine phosphorylation-based regulatory network in *B. subtilis*. In unstressed cells the sigma factor is sequestered by kinase RsbW and anti-anti sigma factor RsbV is in a phosphorylated state unable to complex RsbW. In response to energy or environmental stress RsbV is dephosphorylated and binds RsbW while sigma factor B is released and activates gene expression (Yang *et al*., 1996). Two stress sensing pathways converge on RsbV. In response to energy stress RsbQ interacts with and activates phosphatase RsbP that in turn dephosphorylates RsbV (Brody *et al*., 2001). Environmental stress is sensed by the 1.8 MD stressosome composed of RsbS and RsbR paralogs A-D that in unstressed cells sequesters kinase RsbT (Chen *et al*., 2003). Upon stress induction RsbT phosphorylates RsbR and RsbS releasing RsbT to interact with and activate phosphatase RsbU that subsequently dephosphorylates RsbV. Finally, phosphatase RsbX, whose expression is controlled by sigma B, dephosphorylates RsbR and RsbS returning the system to the unstressed state (Kim *et al*., 2004). A similar module of partner switching is found in the developmental process sporulation in *B. subtilis* where kinase and anti-sigma factor SpoIIAB is counteracted by anti-anti sigma factor SpoIIAA. SpoIIAB binding specificity is strongly dependent on ATP/ADP ratio. ATP stimulates formation of SpoIIAB-sigma factor F complex and further stimulates inactivation of SpoIIAA by phosphorylation while ADP favours SpoIIAB-SpoIIAA complex (Alper *et al*., 1994). Serine phosphatase SpoIIE dephosphorylates SpoIIAA thereby promoting SpoIIAA-AB complexing and release of sigma F (Duncan *et al*., 1995).

Hanks type kinases have received somewhat less attention with PrkC being the best characterised member. PrkC is co-transcribed with its cognate phosphatase PrpC and an early study demonstrated that deletion of either kinase or phosphatase affects sporulation efficiency and reduces biofilm formation (Madec *et al*., 2002). The role in sporulation was further highlighted by a study that

demonstrated an alternative spore germination pathway proceeding via PrkC. Here, the PASTA domain of transmembrane kinase PrkC binds muropeptides activating PrkC and ultimately leading to germination (Shah *et al*., 2008). Muropeptides are also an activating signal in growing cells where a PrkC substrate EF-G is phosphorylated in response to muropeptides. Further, a secreted murolytic enzyme YocH is expressed in a PrkC-dependent manner (Shah and Dworkin, 2010). Underscoring the global importance of the PrkC/PrpC system, it has been shown to control the phosphorylation status of EF-G, EF-TU and CpgA involved in protein synthesis, the stress protein YezB as well as YwjH, GlnA, Icd and AlsD in central metabolism and the important regulator protein HPr (Gaidenko *et al*., 2002; Absalon *et al*., 2009; Pietack *et al*., 2010).

Tyrosine phosphorylation in *B. subtilis* physiology

At present, four proteins targeted by BY-kinase PtkA and its cognate phosphatase PtpZ have been identified. These substrates are the UDP-glucose dehydrogenases Ugd and TuaD and single-stranded DNA binding proteins SsbA and SsbB (YwpH) (Mijakovic *et al*., 2003; Mijakovic *et al*., 2006).

Phosphorylation of the UDP-glucose dehydrogenases Ugd and TuaD increases their activity thereby stimulating production of glucuronic acid, a precursor for production of the exopolysaccharide teichuronic acid (Mijakovic *et al*., 2003). In phosphate-limiting conditions the phosphate-rich teichoic acid of the cell wall is partially substituted with the non-phosphate containing polymer teichuronic acid thereby freeing phosphate for other cellular purposes. Under these conditions the synthesis of glucuronic acid is maintained by TuaD (Pagni *et al*., 1999). The biological relevance of TuaD phosphorylation is still to be confirmed, since an analysis of wild type and Δ*ptkA* strains failed to show a difference in the uronic acid pool in phosphate starved conditions (Petranovic *et al*., 2007). In phosphate-replete conditions where Ugd would be the relevant enzyme, exponentially growing Δ*ptkA* and Δ*ugd* strains showed a 30-40 % decrease in the uronic acid pool as compared to wild type cells. The role of maintaining the glucuronic acid pool under these conditions is not clear (Petranovic *et al*., 2007).

Phosphorylation of the single-stranded DNA-binding proteins SsbA and SsbB affects their single-stranded DNA-binding activity *in vitro* and therefore might affect different stages of DNA metabolism (Mijakovic *et al*., 2006). SsbB is involved in natural competence (Hahn *et al*., 2005),

but natural competence is not significantly affected in Δ*ptkA* and Δ*ptpZ* strains indicating that SsbB phosphorylation may not be relevant in competence (Petranovic *et al*., 2007). SsbA is involved in DNA replication and is essential in *B. subtilis* (Lindner *et al*., 2004; Kobayashi *et al*., 2003). A Δ*ptkA* strain shows a strong phenotype in DNA replication and cell cycle control indicating that SsbA phosphorylation is important in these processes (Petranovic *et al*., 2007).

The published *B. subtilis* phosphoproteome identified an additional nine proteins phosphorylated on tyrosine including several enzymes in carbon metabolism as well as proteins involved in peptide transport, protein synthesis, DNA metabolism and motility indicating a more prominent role of tyrosine phosphorylation in *B. subtilis* physiology.

Bacillus subtilis – a model for multicellular and social behaviour

The main body of *B. subtilis* research has been done on the tryptophan auxotrophic strain 168 originally isolated by Burkholder and Giles after UV mutagenesis using planctonic cultures. A study by Branda and co-workers (Branda *et al*., 2001) however opened for another level of complexity to *B. subtilis* by demonstrating that natural isolates, opposed to laboratory strains, formed robust, highly structured pellicles (biofilm in the liquid/air interface). Further they demonstrated that colonies of natural isolates formed aerial structures that functioned as the preferred site of sporulation. This already brought the challenging view that sporulation could be a process essentially assuring the dispersal of cells rather than or in complement to being a process devoted survival under starvation. Other studies followed up on these findings and have in more detail studied multicellular traits such as pellicle formation and swarming where the two component system DegS/U was shown to play important roles (Kearns *et al*., 2004; Verhamme *et al*., 2007; Kobayashi *et al*., 2007). DegS/U is particular in that the histidine kinase DegS is cytosolic and therefore likely to sense intracellular signals, and that the response regulator DegU is active in both its non-phosphorylated and phosphorylated state (Msadek *et al*., 1990). The importance of DegS/U is underscored by the complex layers of regulation exerted on the system. DegS was recently found to be phosphorylated on a serine residue located in its sensing domain (Murray *et al*., 2009; Macek *et al*., 2007). In this PhD work the role of serine phosphorylation of DegS was studied (chapter 3).

1.6 Outline of the thesis work

The site specific phosphoproteome of *B. subtilis* identified many new phosphoproteins and in most cases identified the phosphorylated residue. This laid the foundation for this thesis work in which I have begun the important next step: the characterisation of the identified phosphorylation events. This resulted in two studies, the first one covering the nine tyrosine phosphorylated proteins (chapter 2) and in the second I more detailed study the serine phosphorylation of the important transition phase two component system kinase DegS (chapter 3).

As a part of Ivan Mijakovic's group I participated in the writing of three review papers dealing with bacterial phosphoproteomics, phosphorylation in relation to pathogenicity and BY-kinases respectively (Soufi *et al*., 2008a; Jers *et al*., 2009; Shi *et al*., 2010) as well as aided in the experimental verification of a bacteria-specific phosphopredictor (Miller *et al*., 2008). Although the phosphopredictor study is related to my project area, my contribution was minimal and as a result I have not included it in this thesis. The topics described in the reviews are covered to some extent in the general introduction. The papers can be found in appendices I-IV.

Tyrosine phosphorylation

With PtkA being the only characterised tyrosine kinase in *B. subtilis*, we asked whether it would be able to phosphorylate any of the nine new tyrosine phosphorylated proteins, and we found that the majority of these proteins could be phosphorylated by PtkA *in vitro*. The previously characterised PtkA substrates all demonstrated an increased activity upon phosphorylation, and we therefore tested whether this would also be the case for the newly identified PtkA substrates. Two of the proteins single-stranded DNA exonuclease YorK and aspartate semialdehyde dehydrogenase Asd were activated by phosphorylation. Since enzyme activity was in many cases not affected we used fluorescent protein fusions to study localisation of these proteins as well as PtkA and its modulator TkmA *in vivo*. For several substrates co-localisation with PtkA was observed, and more importantly, the localisation pattern of the proteins enolase, YjoA, YnfE, YvyG, Ugd and SsbA was dramatically altered in Δ*ptkA* background. Further the localisation data pointed towards a role for the domain organisation in *Firmicutes* where kinase and modulator domains are found in two separate polypeptides.

The most important finding was that PtkA besides its previously described role of controlling enzyme activity of its substrates, also seemed to ensure correct cellular localisation of its targets, which is a new mode of action not attributed to BY-kinases before.

Serine phosphorylation of DegS

The two component system DegS/U has been extensively studied for more than 25 years and still new aspects of this intriguing system are discovered. DegS/U is an important regulatory system that influences many of the transition phase phenomena such as competence, swarming, complex colony and biofilm formation as well as exoprotease production. The regulation of DegS/U is extremely complex and takes place both on transcriptional and protein level. The phosphorylation level of response regulator DegU is fine-tuned and triggers different sub-regulons at different levels. Two component system sensory kinases typically sense a signal via their N-terminal signal sensing domain, but no such signal has been reported for DegS. In the phosphoproteome study, DegS was found to be phosphorylated on serine 76 located in its sensing domain, and we therefore hypothesised that phosphorylation could be an activating signal integrated by DegS/U.

In this study we identified two Hanks type kinases, YbdM and YabT, able to phosphorylate DegS *in vitro* and this stimulated phosphate transfer towards DegU. Next we examined the consequences *in vivo* using strains in which the wild type *degS* gene was replaced with versions encoding a phosphomimetic mutant DegS S76D and a non-phosphorylatable mutant DegS S76A. In different physiological assays the DegS S76D mutant behaved like a strain with intermediate levels of DegU phosphorylation while DegS S76A behaved like a strain with low levels of DegU phosphorylation suggesting a link between DegS phosphorylation at serine 76 and the level of DegU phosphorylation.

This study established serine phosphorylation of DegS as an additional trigger for this two-component system and further represents the first example of two component system sensory kinase regulated by phosphorylation of its signal sensing domain by a Hanks type serine/threonine kinase.

Chapter 2

Tyrosine phosphorylation in *Bacillus subtilis*

Bacillus subtilis BY-kinase PtkA controls enzyme activity and localisation of its protein substrates

[1,2]Carsten Jers, [2]Malene Mejer Pedersen, [2]Dafni Katerina Paspaliari, [4]Wolfgang Schütz, [2]Christina Johnsson, [2,3]Boumediene Soufi, [4]Boris Macek, [2]Peter Ruhdal Jensen and [1,2]Ivan Mijakovic*

[1] Microbiologie et Génétique Moléculaire, AgroParisTech-INRA-CNRS, FR-78850 Thiverval-Grignon, France

[2] Center for Systems Microbiology, Technical University of Denmark, DK-2800 Lyngby, Denmark

[3] Center for Biological Sequence Analysis, Technical University of Denmark, DK-2800 Lyngby, Denmark

[4] Proteome Center Tübingen, University of Tübingen, D-72076 Tübingen, Germany

*For correspondence. E-mail: Ivan.Mijakovic@grignon.inra.fr, Tel: +33(0)1 30 81 45 40, Fax: +33(0)1 30 81 54 57

Accepted for publication in Molecular Microbiology

Summary

Bacillus subtilis BY-kinase PtkA was previously shown to phosphorylate, and thereby regulate the activity of two classes of protein substrates: UDP-glucose dehydrogenases and single-stranded DNA-binding proteins. Our recent phosphoproteome study identified nine new tyrosine phosphorylated proteins in *B. subtilis*. We found that the majority of these proteins could be phosphorylated by PtkA *in vitro*. Among these new substrates, single-stranded DNA exonuclease YorK, and aspartate semialdehyde dehydrogenase Asd were activated by PtkA-dependent phosphorylation. Since enzyme activity was not affected in other cases, we used fluorescent protein tags to study the impact of PtkA on localisation of these proteins *in vivo*. For several substrates co-localisation with PtkA was observed, and more importantly, the localisation pattern of the proteins enolase, YjoA, YnfE, YvyG, Ugd and SsbA was dramatically altered in Δ*ptkA* background. Our results confirm that PtkA can control enzyme activity of its substrates in some cases, but also reveal a new mode of action for PtkA, namely ensuring correct cellular localisation of its targets.

Introduction

Bacterial tyrosine kinases (BY-kinases) are autophosphorylating bacterial enzymes, with no orthologues in *Eukarya* that are involved in many facets of cellular regulation (Grangeasse *et al*., 2007). BY-kinases autophosphorylate on a number of tyrosine residues in a C-terminal tyrosine cluster, and more notably, they have been found to control a number of key enzyme activities by means of phosphorylating their protein substrates (Mijakovic *et al*., 2003; Klein *et al*., 2003; Mijakovic *et al*., 2006; LaCour *et al*., 2008; Minic *et al*., 2007). *In vivo*, BY-kinase activities are countered by those of cognate protein-tyrosine phosphatases that dephosphorylate BY-kinases and their substrates (Vincent *et al*., 1999; Mijakovic *et al*., 2005a). Among several roles of BY-kinases, the best characterised is their involvement in production of extracellular polysaccharides, where they directly participate in export and polymerisation of sugar units (dependent on kinase autophosphorylation) and also regulate some key enzymes by means of phosphorylation (Whitfield, 2006). *Escherichia coli* Wzc (Vincent *et al*., 1999; Grangeasse *et al*., 2003) and PtkA from *Bacillus subtilis* (Mijakovic *et al*., 2003) are arguably the two most extensively characterised BY-kinases from the biochemical perspective. These enzymes represent two different BY-kinase architectures (Grangeasse *et al*., 2007), characteristic respectively for *Proteobacteria* and *Firmicutes*, with the former consisting of a single membrane-spanning polypeptide and the latter being split into a cytosolic kinase and a transmembrane activator. PtkA is thus a soluble cytosolic protein, activated by a specific interaction with its transmembrane modulator TkmA (Mijakovic *et al*., 2003). Once activated, PtkA was found to phosphorylate a number of protein substrates that belong to two major classes: UDP-glucose dehydrogenases (Ugd (old name YwqF) and TuaD) (Mijakovic *et al*., 2003; Mijakovic *et al*., 2004; Petranovic *et al*., 2009) and single-stranded DNA-binding proteins (SsbA and SsbB (old name YwpH)) (Mijakovic *et al*., 2006; Petranovic *et al*., 2007). Both classes of enzymes were found to be activated by PtkA-dependent phosphorylation, and inactivated upon dephosphorylation by the cognate phosphatase PtpZ (Mijakovic *et al*., 2005b). Historically, most bacterial protein kinases were found to be specific for one protein substrate, therefore substrate promiscuity of PtkA was initially surprising to some degree. However, taking into account that eukaryal protein kinases usually phosphorylate a large number of proteins (Ptacek *et al*., 2005), and the fact that the number of phosphoproteins in bacterial phosphoproteomes (Macek *et al*., 2007; Macek *et al*., 2008; Soufi *et al*., 2008b) surpasses the number of known kinases by a factor of 10, relaxed kinase specificity is not entirely unexpected. Recently the first structure of a BY-kinase, a

PtkA homologue CapB from *Staphylococcus aureus*, was resolved (Olivares-Illana *et al*., 2008). This *Firmicute*-type BY-kinase is an octamer anchored to the membrane via its interaction with the transmembrane modulator CapA (analogue of *B. subtilis* TkmA). Upon autophosphorylation, CapA/B octamer was suggested to dissociate, with each CapB monomer staying attached to its corresponding CapA partner. Assuming that PtkA is also found in an octameric complex, we speculated that PtkA, unlike CapB, might dissociate from its modulator TkmA under certain conditions, which would allow it to "seek out" cytosolic substrates and potentially also other modulators. The recently published phosphoproteome of *B.* subtilis (Macek *et al*., 2007) reported the existence of nine new tyrosine-phosphorylated proteins, and we set out to examine whether PtkA might phosphorylate any of them. We have thus documented a number of new proteins phosphorylated by PtkA *in vitro*, two of which were activated by phosphorylation. More interestingly, a number of the new PtkA substrates changed their cellular localisation in a PtkA-dependant manner, indicating for the first time that tyrosine phosphorylation can influence protein targeting in addition to modulating enzyme activities in bacteria.

Results

PtkA phosphorylates a number of new substrates *in vitro*

In order to examine whether any of the nine tyrosine-phosphorylated proteins identified in *B. subtilis* by site-specific phosphoproteomics (Macek *et al*., 2007) could be substrates of PtkA, we purified them as 6xHis-tagged fusions by affinity chromatography. We then performed *in vitro* phosphorylation assays, incubating these proteins with PtkA and its adapter TkmA-NCter in the presence of ^{32}P-γ-ATP. As shown previously (Mijakovic *et al*., 2003), PtkA autophosphorylates weakly even in the absence of TkmA (Figure 2.1, lanes 1). Under the employed experimental conditions none of the substrates auto-phosphorylated (data not shown), and none were phosphorylated by PtkA in the absence of the transmembrane modulator TkmA (Figure 2.1, lanes 1). All nine proteins were phosphorylated by PtkA in the presence of its modulator, with various efficiencies (Figure 2.1, lanes 3). YnfE was found to be heavily phosphorylated, while proteins Asd, InfA, YjoA, YorK and YvyG were moderately phosphorylated and Ldh and OppA were weakly phosphorylated. Enolase incorporated a very low level of radioactivity, making it unclear whether it could indeed be considered a substrate of PtkA. In order to confirm these finding, and verify

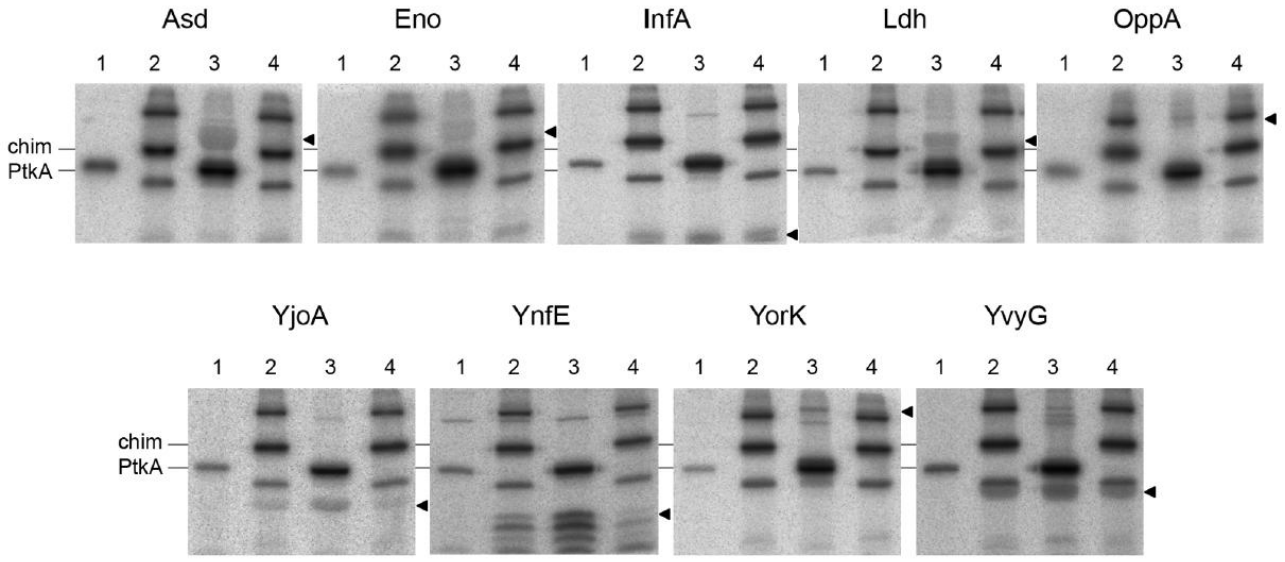


Figure 2.1: *In vitro* phosphorylation assays. Purified protein substrates were incubated with either PtkA alone (lanes 1), chimeric kinase (lanes 2), PtkA and TkmA-NCter (lanes 3) or chimeric kinase with TkmA NCter (lanes 4). Proteins were incubated with ^{32}P-γ-ATP and 5 mM $MgCl_2$ for 1 hour, separated by SDS-PAGE and visualized by STORM PhosphorImager (GE Healthcare). For each gel the positions of PtkA and the chimeric kinase (chim) are indicated with a line, and the protein substrate position is indicated with an arrow to the right of each gel.

whether the phosphorylation sites correspond to the ones detected *in vivo* (Macek *et al*., 2007), we performed a mass spectrometry analysis of non-radioactive *in vitro* phosphorylation mixtures. All substrate proteins were treated by the alkaline phosphatase prior to PtkA-dependent phosphorylation, to remove any potential phosphorylation originating during protein synthesis in *E. coli*. All nine proteins were phosphorylated on tyrosine and in case of InfA, Ldh, OppA, YjoA, YorK and YvyG the previously reported phosphorylation sites were confirmed (Table 2.1). Interestingly, we also identified a number of new tyrosine-phosphorylation sites, whose physiological relevance remains questionable since they have only been identified *in vitro*. Our next question was the importance of the dissociation capability of the kinase PtkA from its adapter TkmA for substrate phosphorylation. We constructed a chimeric kinase resembling BY-kinases of *Proteobacteria*, by fusing the C-terminal part of the modulator TkmA to the N-terminus of PtkA and tested its phosphorylation properties *in vitro*. Phosphorylation assays demonstrated autophosphorylation of the chimeric kinase to be equivalent to that of wild type PtkA activated by its modulator (Figure 2.1, lanes 2). As expected, the addition of TkmA-NCter could not further stimulate the activity of the chimeric kinase (Figure 2.1, lanes 4), since the PtkA region suspected to interact with TkmA was blocked by the fusion with the C-terminal part of TkmA. However, the ability of the chimeric kinase to phosphorylate its substrates was diminished compared to the wild

Table 2.1: Phosphorylation sites on substrates phosphorylated *in vitro* by PtkA, determined by mass spectrometry. Only unambiguously determined sites are shown.

Protein	Previously identified sites[1]	Confirmed	New sites from this study
InfA	Y60	YES	none
Ldh	Y224	YES	Y66, Y69, Y207, Y235
OppA	Y301 or Y303	YES (both)	Y123, Y212, Y395, Y538
YjoA	Y150	YES	Y30, Y32
Asd	Y146	NO	Y289
YnfE	Y12	NO	Y24, Y53, Y55
YorK	Y473	YES	Y3, Y11, Y168, Y220, Y368, Y473
Eno	Y281	NO	Y8, Y46, Y249, Y256, Y403, Y419, Y424, Y426
YvyG	Y49	YES	none

[1]From Macek *et al*., 2007

type PtkA. It was capable of phosphorylating InfA, YjoA, YnfE and YvyG to some extent, but phosphorylation of the other substrates was not detected (Figure 2.1, lanes 2 and 4). Overproduction and purification of the chimeric kinase from *E. coli* resulted in a low yield and co-purification of two proteins that were also strongly phosphorylated *in vitro*. Whether these could be e.g. dimers of the chimeric kinase and degradation products remains to be established.

Asd and YorK are activated by PtkA-dependent phosphorylation *in vitro*

Having demonstrated that PtkA could, at least to some extent, phosphorylate all nine new tyrosine-phosphorylated proteins, we wanted to test whether phosphorylation would influence their enzyme activities, as demonstrated for the previously characterised PtkA substrates (Mijakovic *et al*., 2003; Mijakovic *et al*., 2006). To this end, *in vitro* assays to test the primary function of the proteins were set up and activities of unphosphorylated and phosphorylated proteins were assayed. In several cases no impact of PtkA-dependent phosphorylation on enzyme activity could be demonstrated. Such was the case of the enzymes involved in the central carbon metabolism, Ldh and enolase, where we performed simple colorimetric assays. For YvyG, a putative flagellar chaperone (Pallen *et al*., 2005), we assayed its ATPase activity as an indirect measure of chaperone activity. OppA is a component of the only tripeptide uptake system in *B. subtilis* (Koide and Hoch, 1994) and, as expected, a Δ*oppA* strain did not take up tripeptides. We tested peptide uptake *in vivo* in wild type and Δ*ptkA* cells but no effect on peptide uptake was observed. In each case we could detect basal enzyme activity, but no effect of phosphorylation was detected. Quantitative overview of

experimental data for non-activated substrates is given in Table 2.S2. In the case of proteins of unknown function, YjoA and YnfE, and the translation initiation factor InfA no activity experiments were conducted. Finally, effects of PtkA-dependent phosphorylation were observed with Asd and YorK. Asd converts aspartyl phosphate to aspartyl semialdehyde and inorganic phosphate, with concomitant oxidation of NADPH. We compared the processivity of the PtkA-phosphorylated enzyme and non-phosphorylated enzyme and demonstrated a more than 3-fold activation of phosphorylated Asd under the conditions employed (Figure 2.2). YorK is annotated as a putative ssDNA-specific exonuclease. We therefore initially tested if YorK would exhibit exonuclease activity in the presence of different divalent cations. The enzyme demonstrated ssDNA exonuclease activity in the presence of Mg^{2+} and Mn^{2+} but not Ca^{2+} and Zn^{2+} (data not shown). We then tested whether phosphorylation of YorK would affect its activity, and here it became apparent that PtkA mediated phosphorylation activates this otherwise rather inefficient enzyme in our experimental conditions (Figure 2.3). No DNA degradation products were observed on the gel, indicating that YorK acted in a processive manner, degrading the entire oligonucleotide. Control reactions confirmed that PtkA alone exhibited no exonuclease activity (Figure 2.S1).

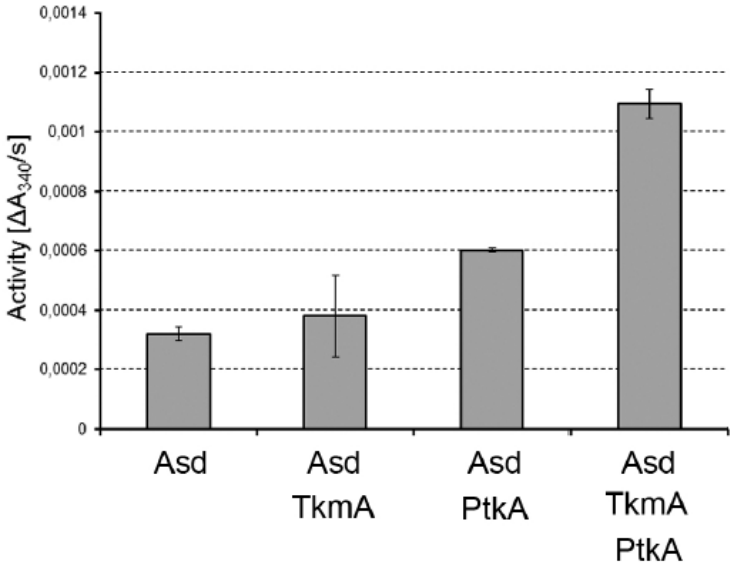


Figure 2.2: Activation of Asd by PtkA-dependent phosphorylation *in vitro*. Asd was incubated alone, with modulator TkmA, with kinase PtkA or with TkmA and PtkA for 2 hours in the presence of 5 mM ATP. Asd reaction was measured as the oxidation of NADPH at 340 nm and initial reaction rates were recorded. The results represent the average of two independent measurements and standard deviations are indicated with error bars.

Cellular localisation of PtkA substrates changes in Δ*ptkA* background

Since the effect of PtkA on enzyme activity was observed for only two of the nine new substrates, we wondered whether PtkA might have other kinds of regulatory effects on the rest of its substrates. To this end, we devised an *in vivo* system to study the impact of PtkA on cellular localisation of its

substrates in *B. subtilis* using fluorescent protein tags. PtkA, TkmA and all known PtkA substrates were produced as fluorescent protein fusions (with either yellow fluorescent protein YFP or cyan fluorescent protein CFP) in wild type *B. subtilis* and Δ*ptkA* background. Initially we confirmed by microscopy that there was no signal from CFP in the yellow filter and YFP in the cyan filter (data not shown). We then co-expressed *ptkA-cfp* with individual target genes fused with *yfp*, and examined the cells with yellow and cyan filter in exponential and stationary phase. First we examined the transmembrane modulator TkmA, which was found to be associated with the membrane where it localised in patches (Figure 2.4, C and D). Interestingly, PtkA co-localised with the modulator in exponential phase (Figure 2.4, E), whereas it seemed to be found primarily in the cytosol in stationary phase (Figure 2.4, F). Next, the two previously characterised substrates of PtkA, SsbA (Mijakovic *et al*., 2006) and Ugd (Mijakovic *et al*., 2003), were examined. SsbA and Ugd both showed a discrete localisation profile, no co-localisation with PtkA was observed, but the localisation pattern changed in a Δ*ptkA* background. SsbA was located in multiple foci in wild-type cells, with the number of foci higher in rapidly dividing cells in exponential phase (Figure 2.5, C and D). In a Δ*ptkA* background, a notable difference was observed in exponential phase, where a sub-population of cells showed a substantially decreased number of foci (Figure 2.5, I). Ugd localised to the pole of wild-type cells in both phases of growth. In the Δ*ptkA* background the protein lost the polar localisation in exponential phase and accumulated in multiple foci in some cells during stationary phase (Figure 2.S2). Next, we examined the new PtkA substrates and concluded that they could be categorised in four groups based on their localisation

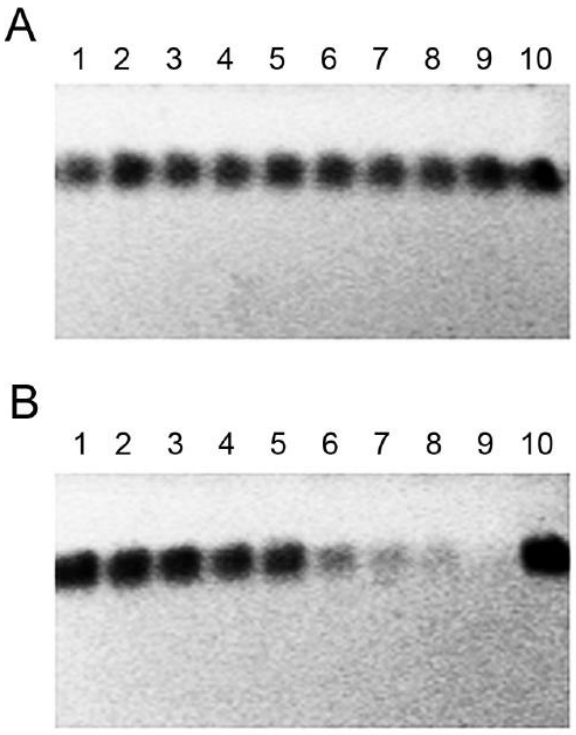


Figure 2.3: Activation of YorK by PtkA-dependent phosphorylation *in vitro*. YorK was pre-incubated either alone (A) or with TkmA and PtkA (B) for 5 hours in the presence of ATP. Exonuclease activity was thereafter measured on an 82-mer oligonucleotide substrate at the following time points: 0.5 h (lane 1), 4 h (lane 2), 6 h (lane 3), 8 h (lane 4), 10 h (lane 5), 14 h (lane 6), 19 h (lane 7), 24 h (lane 8), 48 h (lane 9). Lane 10 contained a control without YorK.

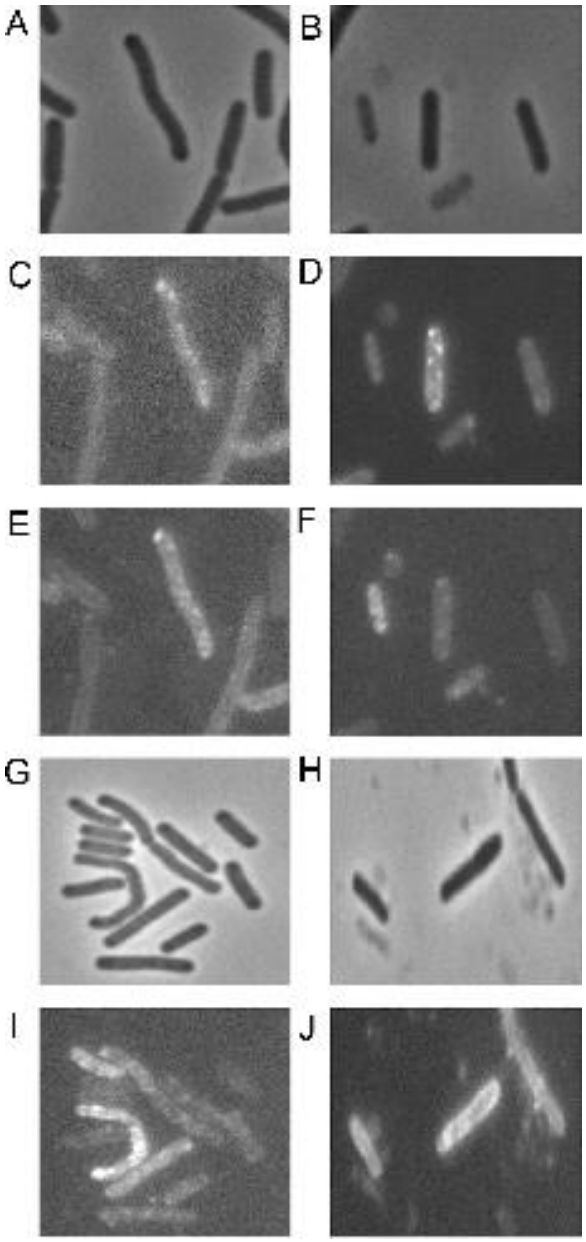


Figure 2.4: Localisation of CFP-PtkA and TkmA-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). TkmA-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). TkmA-YFP visualized in Δ*ptkA* cells in exponential (I) and stationary phase (J).

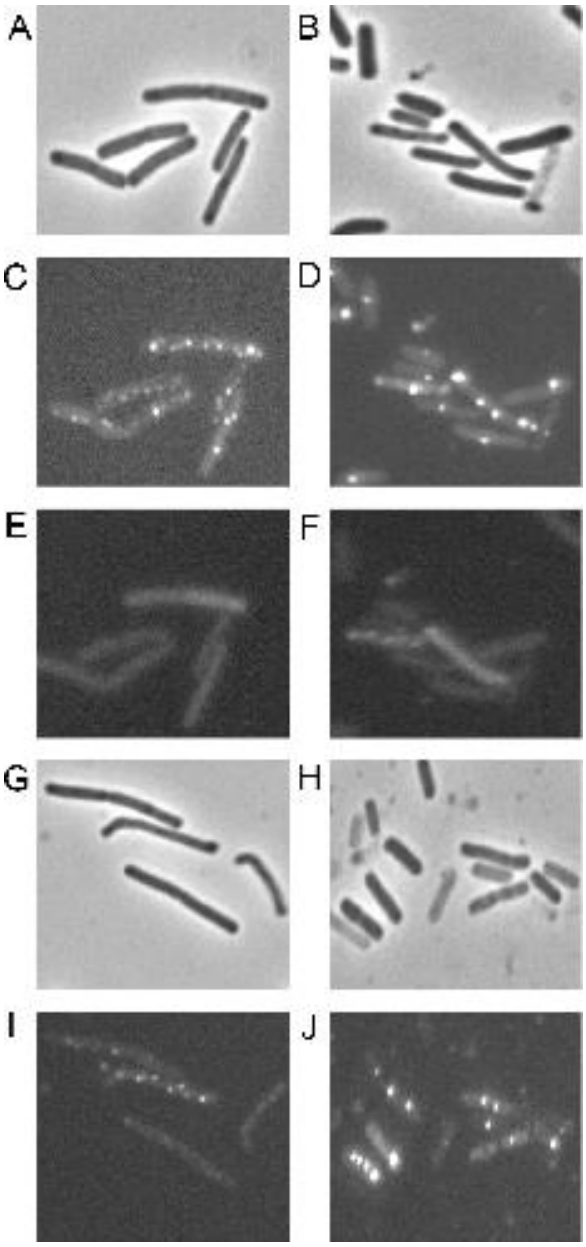


Figure 2.5: Localisation of CFP-PtkA and SsbA-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). SsbA-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). SsbA-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

profiles. The proteins Asd, YorK and InfA exhibited a diffuse localisation profile in both growth phases (Figures 2.S3-S5, respectively) and were not affected by *ptkA* knockout. Enolase (Figure 2.6) and YjoA (Figure 2.S6) exhibited localisation profiles that varied according to the growth

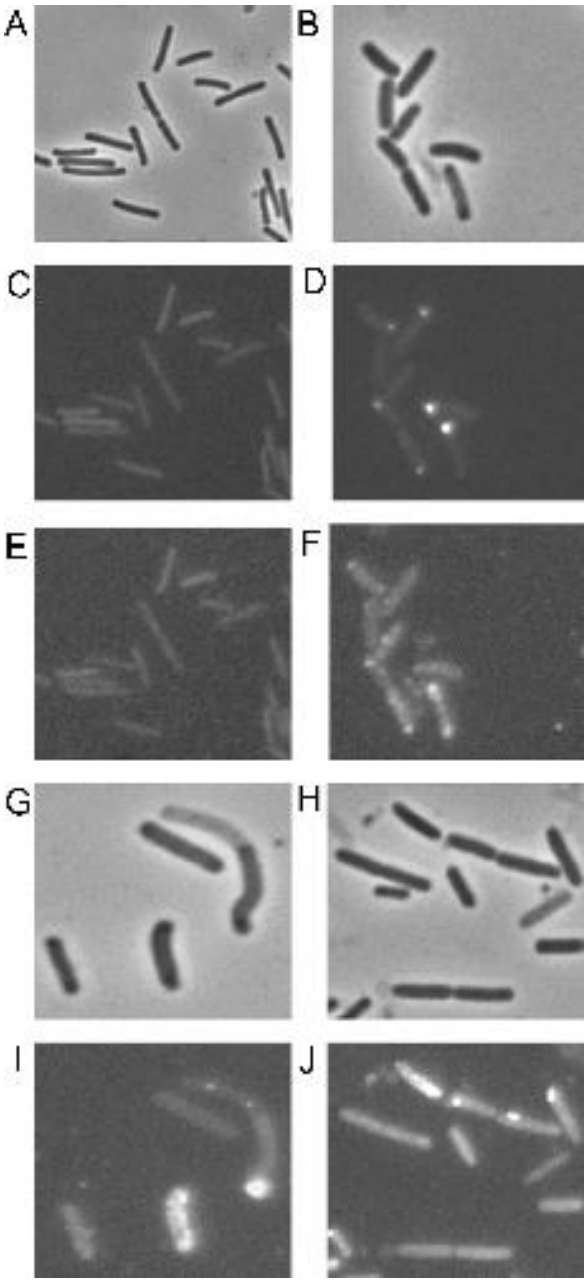


Figure 2.6: Localisation of CFP-PtkA and Eno-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). Eno-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). Eno-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

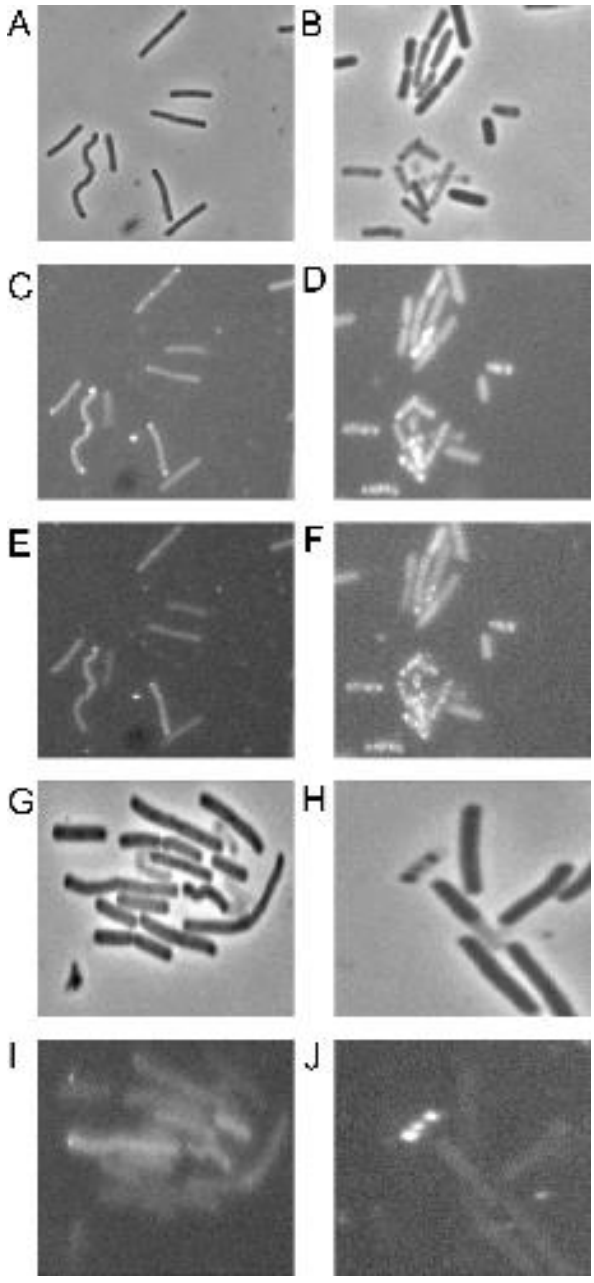


Figure 2.7: Localisation of CFP-PtkA and YvyG-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). YvyG-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). YvyG-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

phase, and were PtkA-dependent, but did not co-localise with PtkA. Enolase exhibited a diffuse profile in the exponential phase (Figure 2.6, C) but localised very strongly to one pole in stationary phase cells (Figure 2.6, D). In Δ*ptkA* cells, enolase foci either disappeared entirely, accompanied by

cytosolic distribution of the enzyme, or appeared as aberrant foci, deformed and with sub-polar localisation (Figure 2.6, J). YjoA was localised in grains in a sub-population of cells during exponential phase and was uniformly distributed in the stationary phase (Figure 2.S6, C and D).

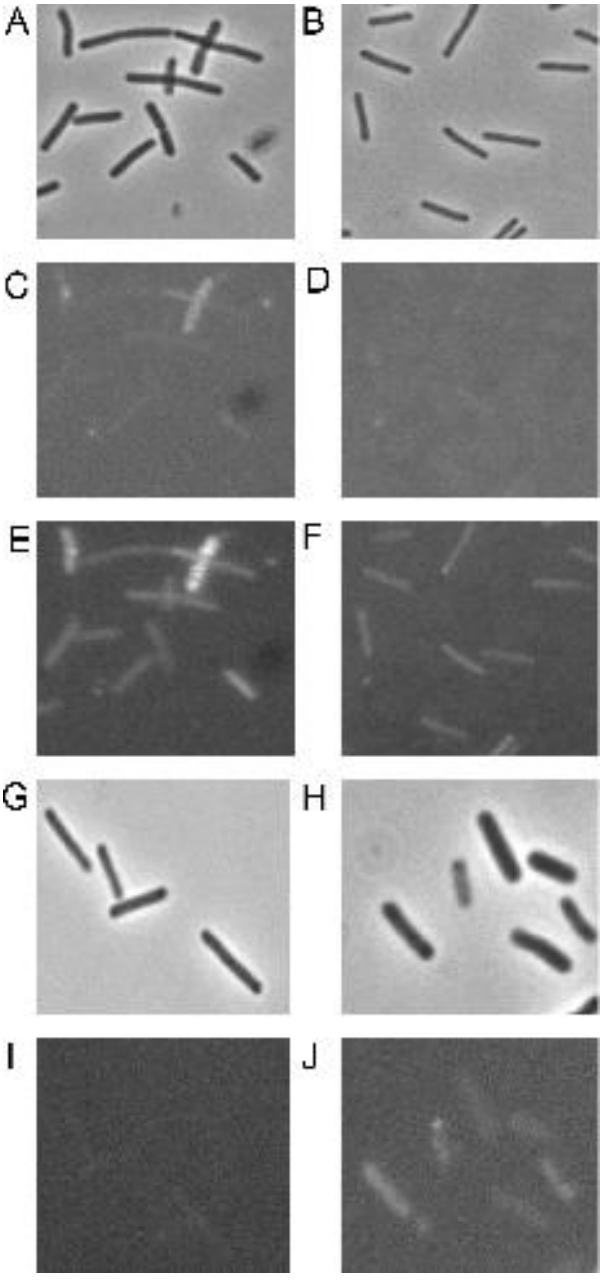


Figure 2.8: Localisation of CFP-PtkA and OppA-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). OppA-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). OppA-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

In the *ptkA* mutant there was no change in exponential phase, while in stationary phase YjoA localised in grains in a sub-population of cells (Figure 2.S6, J). Ldh (Figure 2.S7), YnfE (Figure 2.S8) and YvyG (Figure 2.7) exhibited a PtkA-dependent localisation profile that varied with the growth phase, and further co-localised with PtkA in stationary but not exponential growth phase. YvyG in particular localised to the poles in the exponential phase (Figure 2.7, C) and in several grains spread out in the cell in stationary phase (Figure 2.7, D), while some cells showed a diffuse profile. This localisation profile disappeared in most cells in a Δ*ptkA* background (Figure 2.7, I and J). Finally, in the case of OppA the protein signal was weak and uniformly distributed in the exponential phase (Figure 2.8, C) while it seemed to disappear in the stationary phase (Figure 2.8, D), when OppA is expected to be exported and employed in peptide transport. In a *ptkA* mutant OppA was observed in the cytosol in the stationary phase, which might indicate accumulation due to a defect in its export (Figure 2.8, J). An overview of the results of this localisation screening is summarised in Table 2.2.

Table 2.2: Summary of experimental results. The growth phase where co-localisation was observed is stated in parenthesis. For proteins exhibiting diffuse profiles, co-localisation cannot be determined.

Protein	Phosphorylation by PtkA *in vitro*	Effect on enzyme activity	Co-localisation with PtkA	PtkA dependent localisation
TkmA	–	–	+ (exp. phase)	–
Asd	++	+	diffuse	–
YorK	++	+	diffuse	–
Enolase	(+)	–	–	+
Ldh	+	–	+ (stat. phase)	+
YjoA	++	*not tested*	–	+
YnfE	+++	*not tested*	+ (stat. phase)	+
YvyG	++	–	+ (stat. phase)	+
InfA	++	*not tested*	diffuse	–
OppA	+	–	NA	NA
Ugd	+++[1]	+[1]	–	+
SsbA	++[2]	+[2]	–	+

[1] Mijakovic *et al*., 2003
[2] Mijakovic *et al*., 2006

Discussion

Unlike single polypeptide chain-membrane spanning BY-kinases in *Proteobacteria*, *Firmicutes* possess BY-kinases that are separated into two polypeptides: a transmembrane modulator and a cytosolic kinase. The solved structure of *S. aureus* BY-kinase CapB (Olivares-Illana *et al*., 2008) revealed an octameric BY-kinase ring structure anchored to the membrane *via* interaction with the transmembrane modulators. Upon autophosphorylation, the BY-kinase octamer dissociates, however, the structural data suggested that kinase monomers are likely to remain associated with monomers of the transmembrane modulator. If the BY-kinase is constantly associated to its modulator, why would *Firmicutes* evolve a split polypeptide chain? Data from our study are beginning to shed some light on this question. We constructed a *Proteobacteria*-like kinase, a chimera consisting of the C-terminus of the modulator TkmA fused to the N-terminus of PtkA. This chimeric kinase was capable of autophosphorylating to similar levels as the wild type PtkA in interaction with TkmA, however, its capacity to phosphorylate protein substrates was severely

impaired. This might suggest that the conformation of the artificial fusion kinase was not flexible enough to accommodate various substrates, but flexibility itself is unlikely to be the reason for splitting the *Proteobacteria*-like kinase in two. Interestingly, our localisation data showed that PtkA co-localised with TkmA at the membrane in the exponential growth phase as expected, but not in the stationary phase, where it was released entirely in the cytosol. Our data suggest that PtkA plays a particularly decisive role in ensuring correct localisation of several of its substrates in the stationary phase, where PtkA also co-localised with some of them. It is presently unclear whether PtkA influences substrate localisation via phosphorylation, protein-protein interaction or other less-direct means. However, these results clearly suggest a new model whereby PtkA can dissociate from the transmembrane modulator under appropriate conditions, thus freeing itself to phosphorylate or otherwise interact with its cytosolic substrates. Our PtkA interactome data obtained with yeast two-hybrid also support this hypothesis, suggesting that PtkA could have soluble activity modulators other than TkmA, related to its particular functions in the cytosol (M.–F. Noirot-Gros and I. Mijakovic, unpublished results). A particularly pertinent question for further research is the identity of signal(s) that control the activity and the interaction pattern of PtkA.

We have previously shown that PtkA phosphorylates two classes of substrates, UDP-glucose dehydrogenases and single-stranded DNA-binding proteins, and plays important roles in DNA replication and cell cycle control (Mijakovic *et al.*, 2003; Mijakovic *et al.*, 2006; Petranovic *et al.*, 2007). These findings were further supported by our localisation data, in particular pertaining to SsbA. SsbA was localised in multiple foci in the exponential phase cells. The overall number of foci we observed was larger than in previous studies (Berkmen and Grossman, 2006; Meile *et al.*, 2006). This might be due to a higher level of *ssbA* expression in our system, which could be classified as a technical issue. Nevertheless, the number of SsbA foci decreased dramatically in PtkA-deficient cells. Assuming that foci ensue from SsbA binding to single-stranded regions of the chromosome, this finding concurs with the reduction of SsbA DNA-binding affinity in Δ*ptkA* background, as observed in our previous studies (Mijakovic *et al*., 2006, Petranovic *et al*., 2007).

Besides PtkA, *B. subtilis* possesses a truncated BY-kinase EpsB which is most likely non-functional (Mijakovic *et al*., 2003) and a kinase McsB that was initially reported as a tyrosine kinase, but has recently been recognised as an arginine kinase (Fuhrmann *et al*., 2009). With a relative shortage of tyrosine kinases, and the identification of nine new proteins phosphorylated on tyrosine (Macek *et al*., 2007), we decided to test whether PtkA would phosphorylate any of them. Here we demonstrated that PtkA could phosphorylate all nine proteins to various degrees *in vitro*. In

order to establish the proteins as true substrates of PtkA we wanted to test if these phosphorylation events would have any regulatory role. The previously characterised substrates were shown to be activated by phosphorylation, and in this study we demonstrated that two of the new protein substrates, Asd and YorK, were also activated by PtkA-mediated phosphorylation (or the presence of PtkA) *in vitro*. Asd intervenes in the aspartate biosynthesis pathway, but interestingly its product aspartate semialdehyde is a precursor for diaminopimelic acid, a constituent of bacterial cell wall peptidoglycan, possibly linking this regulatory event to cell wall metabolism. YorK is one of two RecJ homologues in *B. subtilis* (Sutera *et al*., 1999) but its function has not been confirmed experimentally prior to this study. Here we initially confirmed exonuclease activity on ssDNA dependent on the requirement of Mg^{2+} or Mn^{2+} as reported for RecJ in some bacteria. Interestingly, RecJ has been shown to interact with single stranded-DNA binding protein in *Haemophilus influenzae* (Sharma and Rao, 2009), and since both interactants are substrates for PtkA in *B. subtilis*, this phosphorylation event further highlights the importance of PtkA in ssDNA metabolism.

Surprisingly, the activity of most of the newly identified PtkA substrates was not affected by phosphorylation. We reasoned that phosphorylation could instead play a role in mediating protein interactions and we therefore examined their localisation in wild-type and Δ*ptkA* background. This approach allowed us to divide the studied proteins into several classes. The proteins that were affected at the activity level, Asd, YorK (and InfA) showed a diffuse localisation profile that was not affected by PtkA inactivation. By contrast, the majority of the proteins showed a growth-phase dependent and PtkA-dependent localisation profile. Ldh, YnfE and YvyG in addition co-localised with PtkA, and enolase and YjoA did not. In the case of Ldh, YnfE, YvyG and YjoA the functional roles of observed localisation profiles are difficult to interpret, but the very strong localisation of enolase to one pole in the cell could indicate a role in sporulation. The multiply phosphorylated protein enolase has been shown to be multifunctional; besides its initially recognised role in glycolysis, it is implicated in heat shock response (Miller *et al*., 1991), RNA degradation (Commichau *et al*., 2009), DNA replication (Jannière *et al*., 2007; Commichau *et al*., 2009) and is also secreted upon automodification with its substrate 2-PG (Boël *et al*., 2004). It is also one of the most abundant proteins in *Bacilli* spores (DelVecchio *et al*., 2006). In one particular case, that of OppA, the protein signal that was absent in the wild type stationary cells appeared in *ptkA* mutant cells. This protein is part of the Opp permease system specific for peptides of 3-5 amino acids length. OppA is the ligand binding protein, and is attached to the outside of the cell via

a lipid anchor. The fact that it was retained in the cytosol in the *ptkA* mutant could point towards phosphorylation of OppA playing a role in its export. However, the import of peptides *in vivo* was not affected in Δ*ptkA* background, which cautions us against forming any final conclusions in this case.

Previous studies on BY-kinase mediated substrate phosphorylation have mainly studied the effect on protein activity. In *Eukarya* however, several examples of proteins showing a tyrosine phosphorylation-dependent localisation have been reported (Madeo *et al*., 1998; Lukong *et al*., 2005) and the results presented here indicate for the first time that BY-kinases can also act in terms of ensuring correct cellular location of their substrates.

Materials and methods

Bacterial strains and growth conditions

E. coli NM522 was used for plasmid propagation in cloning experiments. The chaperone overproducing strain *E. coli* M15 carrying pREP4-GroESL (Amrein *et al*., 1995) was used for overproduction of protein. *B. subtilis* 168 was used in localisation experiments and *in vivo* assays. *E. coli* and *B. subtilis* strains were grown in LB medium shaking at 37 °C. *B. subtilis* was grown in C-medium (Martin-Verstraete *et al*., 1990) with 5 g/L glucose (CG-medium). When relevant, 100 µg/mL ampicillin, 25 µg/mL kanamycin and 8 µg/mL tetracycline for *E. coli* and 5 µg/mL erythromycin and 15 µg/mL tetracycline for *B. subtilis* were added to the medium.

DNA manipulations and strain construction

Genes *asd, eno, infA, ldh, oppA, yjoA, ynfE, yorK, yvyG* and *yclM* were PCR amplified using *B. subtilis* 168 genomic DNA and specific primers with restriction sites (Table 2.S1) and inserted between the *Bam*HI and *Cfr*9I sites of pQE30-Xa (Qiagen). Vectors for expression of PtkA and TkmA-NCter were described previously (Mijakovic *et al*., 2003). pQE30-strep encoding a Strep-tag in place of a His-tag was constructed by restricting pQE30 with *Bam*HI and *Eco*RI, annealing the 5' phosphorylated oligos pQE-str+ and pQE-str- and ligating the fragments. For production of strep-tagged PtkA, the *Bam*HI-*Pst*I fragment from pQE30-ptkA (Mijakovic *et al*., 2003) was inserted in pQE30-strep. The chimeric kinase gene encoding the 50 C-terminal amino acids of TkmA fused to the N-terminus of PtkA was constructed by PCR using an overlapping primer and inserted between

the *Bam*HI and *Pst*I sites of pQE30. The vector pG$^+$Host8 (Maguin *et al*., 1996) that contains a temperature sensitive origin of replication and a tetracycline resistance cassette was used to construct *B. subtilis* Δ*ptkA* by deletion of the middle part of *ptkA*. The upstream and downstream parts of *ptkA* were amplified with primer pairs ΔptkA up fwd and rev and ΔptkA down fwd and rev respectively. Vector and PCR products were restricted with appropriate enzymes and ligated. This construct was used to transform *B. subtilis* 168 and transformants were incubated at 37 °C (non-replicative temperature) on LB plates containing 15 µg/mL tetracycline for chromosome integration. The integrants were re-streaked and incubated at 28 °C (replicative temperature) without selective pressure to induce excision followed by a shift to 37 °C for loss of the vector. The deletion was confirmed by PCR on genomic DNA. The *oppA* gene was inactivated using the vector pMUTIN2 (Vagner *et al*., 1998). For examination of protein localisation, eYFP (Clontech) and CFP+ were used. Gene *yfp* was PCR amplified using primers yfpC fwd and yfpC rev and a miniTn7-*eyfp* delivery plasmid (Lambertsen *et al*., 2004) as template and inserted between the *Bam*HI and *Cfr*9I sites of pHT315 (Arantes and Lereclus, 1991) to generate pHT315-*yfpC*. In order to avoid possible low CFP signal due to slow translation initiation the first eight codons of *comGA* were encoded on the primer cfpN fwd (Veening *et al*., 2004). *cfp+* (Andersen *et al*., 2006) was amplified using primers cfpN fwd and cfpN rev and inserted between the *KpnI* and *Eco*RI sites of pHT315-*yfpC* to generate pHT315-*yfpC-cfpN*. Gene *ptkA* was amplified using primers pktAN fwd and ptkAN rev and inserted between the *Cfr*9I and *Eco*RI sites of pHT315-*yfpC-cfpN* to generate pHT315-*yfpC-cfp:ptkA*. Genes encoding the phosphotyrosine proteins and TkmA were PCR amplified with relevant primers and inserted between the *Bam*HI and *Avr*II sites of pHT315-*yfpC* and pHT-*yfpC-cfp:ptkA* to generate pHT315-"*gene of interest":yfp* and pHT-"*gene of interest":yfp-cfp:ptkA* respectively. The vector pHT315-*cfp:ptkA* was constructed by restricting pHT315-*yfpC-cfp:ptkA* with *Eco*RI and *Kpn*I and inserting the *cfp:ptkA* fragment in pHT315. *B. subtilis* Δ*ptkA* was transformed with pHT315-*"gene of interest":yfp* and *B. subtilis* 168 was transformed with pHT315-*"gene of interest":yfp-cfp:ptkA* and pHT315-*cfp:ptkA*.

Production and purification of 6xHis-tagged proteins

His-tagged proteins were synthesised in the chaperone overproducing strain *E. coli* M15 carrying pREP4-groESL. Cultures were grown shaking at 37 °C to OD_{600} 0.5, induced with 1 mM IPTG and grown an additional 3 hours. His-tagged proteins were purified on Ni-NTA columns (Qiagen) as described previously (Mijakovic *et al*., 2003), desalted with PD-10 columns and stored in a buffer

containing 50 mM Tris-Cl pH 7.5, 100 mM NaCl and 10 % glycerol. Proteins for mass spectrometry analysis were further dephosphorylated with shrimp alkaline phosphatase (Fermentas). Crude extract from 1 L culture were incubated with 30 units at 37 °C for 1 hour.

In vitro phosphorylation assays

Phosphorylation assays were performed essentially as described previously (Mijakovic *et al*., 2003), with 4.5 nM Ldh, 0.48 nM OppA, 17 nM YjoA, 18 nM YvyG, 26 nM InfA, 45 nM YnfE and 0.17 nM chimeric kinase or 0.23 nM PtkA and TkmA-NCter. For YorK, 2.0 nM, 43 nM and 18 nM Asd and enolase were used respectively, with 0.55 nM chimeric kinase or 0.75 nM PtkA and TkmA-NCter. After 1 hour incubation at 37 °C, proteins were separated by electrophoresis and radioactive signals were visualised with a Storm 860 PhosphorImager. For non-radioactive assays, destined for mass spectrometry analysis, the same procedure was applied but with non-radioactive ATP.

Mass spectrometry analysis of *in vitro* phosphorylation sites

Proteins from *in vitro* phosphorylation reactions (25-100 µg) were diluted with denaturation buffer (6 M urea, 2 M thiourea in 10 mM Tris, pH 7.0) to a final concentration of 1-2 µg/µl, reduced, alkylated and digested with endoproteinase Lys-C and trypsin as described previously (Miller *et al*., 2009). A part (10%) of the sample was taken and the peptides were desalted directly using C18 StageTips (Rappsilber *et al*., 2007). The rest was subjected to phosphopeptide enrichment using TiO_2 beads (GL Sciences) after adding acetonitrile (ACN) to a final concentration of 30%. The beads were preincubated with a solution containing 20 mg/ml dihydrobenzoic acid in 80% ACN and 5 mg of beads were added to each sample. Following the incubation for one hour with end-over-end rotation, the beads were washed twice in a solution containing 60% ACN and 0.1% trifluoroacetic acid (TFA) and eluted with 150µl of 40% ammonia solution in 60% ACN (pH 10.5). Eluates were prepared for LC-MS by reducing their volumes to 5 µl and adding an equal volume of a solution containing 2% ACN and 1% TFA. The peptide mixtures were analyzed using a Proxeon Easy-LC system (Proxeon Biosystems) coupled to a LTQ-Orbitrap-XL mass spectrometer (ThermoFisher) equipped with a nanoelectrospray ion source (Proxeon Biosystems). Chromatographic separation and mass spectrometry were performed essentially as described previously (Miller *et al*., 2009). An inclusion list containing the *m/z* values of peptides carrying previously identified phosphorylation sites (Macek *et al*., 2007) was used. Mass spectra were analysed using the software suite MaxQuant, version 1.0.14.3 (Cox *et al*., 2009). The data were

searched against database of *B. subtilis* (forward primary annotation database was downloaded from http://cmr.jcvi.org) supplemented with tagged versions of used targets and frequently observed contaminants and concatenated with reversed copies of all sequences using MASCOT (version 2.2.0, Matrix Science, London, UK). Carbamidomethylation of cysteine was set as fixed modification and oxidation of methionine, N-terminal acetylation, and phosphorylation (STY) as variable modifications. Initial mass tolerance was set to a maximum of 7 ppm and a maximum of two missed cleavages was allowed. Maximum false discovery rates (FDR) were set to 1% for both, peptide and protein levels. Phosphorylation events were considered to be localized to a specific site if the calculated localisation probability was above 0.75.

Enolase activity assay

Enolase catalyses the conversion of 2-phospho-D-glycerate to phosphoenolpyruvate (2-PGA) and H_2O. To phosphorylate enolase, 0.84 µM enolase was incubated with 1.5 µM PtkA, 1.1 µM TkmA-NCter, 5 mM ATP, 5 mM MgCl, 50 mM Tris-Cl (pH 7.5), 100 mM NaCl and 10 % glycerol for about 16 hours at 37^{o}C. The sample without PtkA and TkmA-NCter was prepared similarly. 25 µL of samples were assayed in total volume of 1 mL comprising 50mM Tris-HCl (pH 7.4), 100 mM KCl, 1 mM $MgSO_4$, 0.01 mM EDTA and 2 mM 2-PGA, as described previously (Boël *et al*., 2004). Reactions were initiated by addition of 2-PGA, followed spectrophotometrically at 240 nm at 37^{o}C, and initial reaction rates were recorded.

Ldh activity assay

Ldh catalyses the conversion of pyruvate and NADH to lactate and NAD^+. $NADP^+$ can be used as co-factor (Romero *et al*., 2007) and Fructose bisphosphate (FBP) has been shown to activate Ldh in some species (Fushinobu *et al*., 1996). Effect of phosphorylation was tested with both NADH and NADPH as co-factor, with and without FBP present. In phosphorylation reactions Ldh was incubated with PtkA, TkmA-NCter, 2.5 mM ATP, 2.5 mM $MgCl_2$, 1 g/L BSA, 50 mM Tris-Cl (pH 7.5), 100 mM NaCl and 10 % glycerol either with or without 0.1 mM FBP for 5 hours at 37 ^{o}C. For NADH experiments 0.56 µM Ldh, 0.47 µM PtkA and 0.47 µM TkmA-NCter and for NADPH experiments were used 8.4 µM Ldh, 7.5 µM PtkA and 7.5 µM TkmA-NCter. Samples without PtkA and TkmA-NCter were prepared similarly. To assay activity, 5.6 nM Ldh for NADH experiments and 84 nM Ldh for NADPH experiments were mixed with 1 mM pyruvate and 1 mM NADH/NADPH with or without 0.1 mM FBP in a buffer containing 50 mM Tris-Cl (pH 7.5), 100

mM NaCl and 10 % glycerol. Reactions were started by addition of pyruvate, followed spectrophotometrically at 340 nm at 37 ^{o}C and initial reaction rates were recorded.

OppA assay

OppA is part of an ABC transporter that imports peptides of 3-5 amino acids length. To test if peptide import was affected by PtkA-mediated phosphorylation we tested sensitivity to the tripeptide antibiotic bialaphos in *B. subtilis* wild type and Δ*ptkA* strains essentially as described previously (Koide and Hoch, 1994). Cultures were grown in CG-minimal medium to OD_{600} of 0.4 (exponential phase) and 1.5 (transition phase). Inhibition zones were measured and results are an average of three experiments.

YvyG chaperone activity assay

YvyG is a putative flagellar chaperone with a putative GTPase domain. Since chaperone activity is ATP-dependent, we attempted to quantify ATP hydrolysis activity as an indirect measure of chaperone activity. Phosphorylation reactions contained 11 µM YvyG, 3.0 µM Strep-tagged PtkA, 2.3 µM TkmA–NCter, 0.5 mM ATP, 5 mM $MgCl_2$, 50 mM Tris-Cl (pH 7.5), 100 mM NaCl and 10 % glycerol. Reactions with no PtkA and TkmA were prepared similarly and reactions were incubated for 3 hours at 37^{o}C. To remove basal ATP hydrolysis activity from PtkA, Strep-tagged PtkA was removed by mixing 95 µL of the phosphorylation reactions with 150 µL Streptactin resin (IBA). ATP hydrolysis was assayed by adding ^{32}P-γ-ATP to a final concentration of 300 µM. Samples were incubated at 37 ^{o}C for 1, 2, 4, and 6 hours, 1 µL was spotted on a polyethylenimine cellulose sheet (Machery-Nagel) and separated by thin layer chromatography using a 0.3 M potassium phosphate buffer (Mijakovic *et al.*, 2002). Radioactive signals were visualised using the STORM PhosphorImager and quantified with ImageQuant (GE Healthcare).

Single-stranded DNA exonuclease assay

To test the effect of phosphorylation on YorK exonuclease activity 0.15 µM YorK was phosphorylated with 0.75 µM PtkA and 0.75 µM TkmA-NCter in a buffer containing 50 mM Tris (pH 7.5), 100 mM NaCl, 8 mM $MgCl_2$, 2 mM ATP and 10 % glycerol for 5 hours. Unphosphorylated YorK was treated in the same buffer without kinase and modulator. The two controls included a phosphorylation reaction without YorK, and one with 115 µM BSA added instead of PtkA. To 12.5 µL phosphorylation reaction, 1,5 µL reaction buffer (20 mM Tris-HCl (pH

8.0), 10 mM MgCl2, 10 mM NaCl, 0.67 mM dithiothreitol (DTT) and 1 mg/mL BSA) and 1 µL of 0.6 pmol of 82-mer oligonucleotide in water. The reactions were incubated at 37 °C for indicated times (see figure legend) and terminated by heat inactivation for 10 minutes at 70 °C. Samples were subjected to gel electrophoresis on 1.5 % agarose gel with ethidium bromide and visualised by UV light.

Asd assay

Asd catalyses the conversion of aspartyl phosphate and NADPH to aspartyl semialdehyde, inorganic phosphate and $NADP^+$. Aspartyl phosphate was produced by incubating 0.99 µM aspartate kinase III (YclM) with 25 mM aspartate-KOH, 10 mM ATP, 100 mM Tris (pH 8), 20 mM $MgCl_2$ and 150 mM KCl for 50 min at 37 °C and subsequently kept on ice. To phosphorylate Asd, 0.78 µM were incubated with 0.50 µM PtkA and 0.75 µM TkmA-NCter in a reaction buffer containing 37.5 mM Tris (pH 7.5), 5 mM $MgCl_2$, 1 mM $MnCl_2$, 5 mM ATP, 10 mg/mL BSA, 75 mM NaCl and 7.5 % glycerol in a total volume of 30 µL for 2 hours at 37 °C. Samples without PtkA, TkmA-NCter or both were prepared similarly. To measure Asd activity, 900 µL Aspartyl phosphate reaction was pre-heated to 37 °C for 5 min before reaction was started by addition of 30 µL Asd preparation and NADPH to a final concentration of 0.5 mM in a 1 mL volume and absorbance at 340 nm was recorded.

Microscopy

For examination of localisation of phospho-tyrosine proteins, TkmA and PtkA, cells were grown in LB at 37 °C with shaking and samples were taken at OD_{600} 0.5 (exponential) and after overnight growth of about 18 h (stationary). Samples were concentrated 10 fold and 5 µL were deposited on a polylysine-coated glass slide (Thermo Scientific) and examined using a Xeiss Axioplan microscope equipped with a Kappa ACC 1 condenser, a Zeiss Plan Neofluor 100x objective and a Kappa DX2 HC-FW camera. Images were acquired using Kappa Imagebase Control 2.7.2 software. For each condition presented in figures, about 200 individual cells were examined and a representative sample was chosen.

Acknowledgements

This work was supported by grants from the Danish National Research Council (FNU), the Lundbeckfonden and the Institut National de Recherche Agronomique (INRA) to IM, the Landesstiftung BW to BM and a PhD stipend from the Technical University of Denmark (DTU) to CJ. We are grateful to Prof. Flemming G. Hansen for help with microscopy, Dr. Sünje Pamp for providing vectors coding for fluorescent proteins, to Peter Boldsen Knudsen and René Jonsgaard Larsen for producing the chimeric kinase during their student project and to Silke Wahl for outstanding technical support.

References

Amrein, K.E., Takacs, B., Stieger, M., Molnos, J., Flint, N.A., and Burn, P. (1995) Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL. *Proc Natl Acad Sci U S A* 92: 1048-1052

Andersen, J.B., Roldgaard, B.B., Lindner, A.B., Christensen, B.B., and Licht, T.R. (2006) Construction of a multiple fluorescence labelling system for use in co-invasion studies of *Listeria monocytogenes*. *BMC Microbiol* 6:86

Arantes, O., and Lereclus, D. (1991) Construction of cloning vectors for *Bacillus thuringensis Gene* 108: 115-119

Berkmen, M.B., and Grossman, A.D. (2006) Spatial and temporal organization of the *Bacillus subtilis* replication cycle. *Mol Microbiol* 62: 57-71

Boël, G., Pichereau, V., Mijakovic, I., Mazé, A., Poncet, S., Gillet, S., *et al*. (2004) Is 2-phosphoglycerate-dependent automodification of bacterial enolases implicated in their export? *J Mol Biol* 337: 485-496

Commichau, F.M., Rothe, F.M., Herzberg, C., Wagner, E., Hellwig, D., Lehnik-Habrink, M., *et al*. (2009) Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. *Mol Cell Proteomics* 8:1350-1360

Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J.V. and Mann, M. (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* 4: 698-705

Delvecchio, V.G., Connolly, J.P., Alefantis, T.G., Walz, A., Quan, M.A., Patra, G., *et al*. (2006) Proteomic profiling and identification of immunodominant spore antigens of *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*. *Appl Environ Microbiol* 72: 6355-6363

Fuhrmann, J., Schmidt, A., Spiess, S., Lehner, A., Turgay, K., Mechtler, K., *et al*. (2009) McsB is a protein arginine kinase that phosphorylates and inhibits the heat-shock regulator CtsR. *Science* 324: 1323-1327

Fushinobu, S., Kamata, K., Iwata, S., Sakai, H., Ohta, T., and Matsuzawa, H. (1996) Allosteric activation of L-lactate dehydrogenase analyzed by hybrid enzymes with effector-sensitive and -insensitive subunits. *J Biol Chem* 271: 25611-25616

Grangeasse, C., Obadia, B., Mijakovic, I., Deutscher, J., Cozzone, A.J., and Doublet, P. (2003) Autophosphorylation of the *Escherichia coli* protein kinase Wzc regulates tyrosine phosphorylation of Ugd, a UDP-glucose dehydrogenase. *J Biol Chem* 278: 39323-39329

Grangeasse, C., Cozzone, A.J., Deutscher, J., and Mijakovic, I. (2007) Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology. *Trends Biochem Sci* 32: 86-94

Jannière, L., Canceill, D., Suski, C., Kanga, S., Dalmais, B., Lestini, R., *et al*. (2007) Genetic evidence for a link between glycolysis and DNA replication. *PLoS One* 2: e447

Klein, G., Dartigalongue, C., and Raina, S. (2003) Phosphorylation-mediated regulation of heat shock response in *Escherichia coli*. *Mol Microbiol* 48: 269-285

Koide, A., and Hoch, J.A. (1994) Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Mol Microbiol* 13: 417-426.

Lacour, S., Bechet, E., Cozzone, A.J., Mijakovic, I., and Grangeasse, C. (2008) Tyrosine phosphorylation of the UDP-glucose dehydrogenase of *Escherichia coli* is at the crossroads of colanic acid synthesis and polymyxin resistance. *PLoS One* 3: e3053

Lambertsen, L., Sternberg, C., and Molin, S. (2004) Mini-Tn7 transposons for site-specific tagging of bacteria with fluorescent proteins. *Environ Microbiol* 6: 726-732

Lukong, K.E., Larocque, D., Tyner, A.L., and Richard, S. (2005) Tyrosine phosphorylation of sam68 by breast tumor kinase regulates intranuclear localization and cell cycle progression. *J Biol Chem* 280: 38639-38647

Macek, B., Mijakovic, I., Olsen, J.V., Gnad, F., Kumar, C., Jensen, P.R., and Mann, M. (2007) The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. *Mol Cell Proteomics* 6: 697-707

Macek, B., Gnad, F., Soufi, B., Kumar, C., Olsen, J.V., Mijakovic, I., and Mann, M. (2008) Phosphoproteome analysis of *E. coli* reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Mol Cell Proteomics* 7: 299-307

Madeo, F., Schlauer, J., Zischka, H., Mecke, D., and Fröhlich, K.U. (1998) Tyrosine phosphorylation regulates cell cycle-dependent nuclear localization of Cdc48p. *Mol Biol Cell* 9: 131-141

Maguin, E., Prévost, H., Ehrlich, S.D., and Gruss, A. (1996) Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. *J Bacteriol* 178: 931-935

Martin-Verstraete, I., Débarbouillé, M., Klier, A., and Rapoport, G. (1990) Levanase Operon of *Bacillus subtilis* Includes a Fructose-specific Phosphotransferase System Regulating the expression of the Operon. *J. mol. biol* 214: 657-671

Meile, J.C., Wu, L.J., Ehrlich, S.D., Errington, J., and Noirot, P. (2006) Systematic localisation of proteins fused to the green fluorescent protein in *Bacillus subtilis*: identification of new proteins at the DNA replication factory. *Proteomics* 6: 2135-2146

Mijakovic, I., Poncet, S., Galinier, A., Monedero, V., Fieulaine, S., Janin, J., *et al*. (2002) Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: A relic of early life? *PNAS* 99: 13442-13447

Mijakovic, I., Poncet, S., Boël, G., Mazé, A., Gillet, S., Jamet, E., *et al*. (2003) Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases. *EMBO J* 22: 4709-4718

Mijakovic, I., Petranovic, D., and Deutscher, J. (2004) How tyrosine phosphorylation affects the UDP-glucose dehydrogenase activity of *Bacillus subtilis* YwqF. *J Mol Microbiol Biotechnol* 8: 19-25

Mijakovic, I., Petranovic, D., Bottini, N., Deutscher, J., and Ruhdal Jensen, P. (2005a) Protein-tyrosine phosphorylation in *Bacillus subtilis*. *J Mol Microbiol Biotechnol* 9: 189-197.

Mijakovic, I., Musumeci, L., Tautz, L., Petranovic, D., Edwards, R.A., Jensen, P.R., *et al*. (2005b) In vitro characterization of the *Bacillus subtilis* protein tyrosine phosphatase YwqE. *J Bacteriol* 187: 3384-3390

Mijakovic, I., Petranovic, D., Macek, B., Cepo, T., Mann, M., Davies, J., *et al*. (2006) Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res* 34: 1588-1596

Miller, B.S., Kennedy, T.E., and Streips, U.N. (1991) Molecular characterization of specific heat shock proteins in *Bacillus subtilis*. *Curr Microbiol* 22: 231-236

Miller, M.L., Soufi, B., Jers, C., Blom, N., Macek, B., Mijakovic, I. (2009) NetPhosBac - Prediction of protein Ser/Thr phosphorylation sites in bacteria. *Proteomics* 9: 116-125

Minic, Z., Marie, C., Delorme, C., Faurie, J.M., Mercier, G., Ehrlich, D., and Renault, P. (2007) Control of EpsE, the phosphoglycosyltransferase initiating exopolysaccharide synthesis in *Streptococcus thermophilus*, by EpsD tyrosine kinase. *J Bacteriol* 189: 1351-1357

Olivares-Illana, V., Meyer, P., Bechet, E., Gueguen-Chaignon, V., Soulat, D., Lazereg-Riquier, S., *et al*. (2008) Structural basis for the regulation mechanism of the tyrosine kinase CapB from *Staphylococcus aureus*. *PLoS Biol* 6: e143

Pallen, M.J., Penn, C.W., and Chaudhuri, R.R. (2005) Bacterial flagellar diversity in the post-genomic era. *Trends Microbiol* 13:143-149.

Petranovic, D., Michelsen, O., Zahradka, K., Silva, C., Petranovic, M., Jensen, P.R., and Mijakovic, I. (2007) *Bacillus subtilis* strain deficient for the protein-tyrosine kinase PtkA exhibits impaired DNA replication. *Mol Microbiol* 63: 1797-1805

Petranovic, D., Grangeasse, C., Macek, B., Abdillatef, M., Gueguen-Chaignon, V., Nessler, S., *et al*. (2009) Activation of *Bacillus subtilis* Ugd by the BY-Kinase PtkA Proceeds via Phosphorylation of Its Residue Tyrosine 70. *J Mol Microbiol Biotechnol* 17: 83-89

Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., *et al* (2005) Global analysis of protein phosphorylation in yeast. *Nature* 438: 679-684

Rappsilber, J., Mann, M. and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2: 1896-1906

Romero, S., Merino, E., Bolívar, F., Gosset, G., and Martinez, A. (2007) Metabolic engineering of *Bacillus subtilis* for ethanol production: lactate dehydrogenase plays a key role in fermentative metabolism. *Appl Environ Microbiol* 73: 5190-5198

Sharma, R., and Rao, D.N. (2009) Orchestration of *Haemophilus influenzae* RecJ exonuclease by interaction with single-stranded DNA-binding protein. *J Mol Biol* 385: 1375-1396

Soufi, B., Gnad, F., Jensen, P.R., Petranovic, D., Mann, M., Mijakovic, I., and Macek, B. (2008) The Ser/Thr/Tyr phosphoproteome of *Lactococcus lactis* IL1403 reveals multiply phosphorylated proteins. *Proteomics* 8: 3486-3493

Sutera, V.A. Jr, Han, E.S., Rajman, L.A., and Lovett, S.T. (1999) Mutational analysis of the RecJ exonuclease of *Escherichia coli*: identification of phosphoesterase motifs. *J Bacteriol* 181: 6098-6102

Vagner, V., Dervyn, E., and Ehrlich, S.D. (1998) A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology*. 144: 3097-3104.

Veening, J.W., Smits, W.K., Hamoen, L.W., Jongbloed, J.D., and Kuipers, O.P. (2004) Visualization of differential gene expression by improved cyan fluorescent protein and yellow fluorescent protein production in *Bacillus subtilis*. *Appl Environ Microbiol* 70: 6809-6815

Vincent, C., Doublet, P., Grangeasse, C., Vaganay, E., Cozzone, A.J., and Duclos, B. (1999) Cells of *Escherichia coli* contain a protein-tyrosine kinase, Wzc, and a phosphotyrosine-protein phosphatase, Wzb. *J Bacteriol* 181: 3472-3477

Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem* 75: 39-68

Supplementary material

Bacillus subtilis BY-kinase PtkA controls enzyme activity and localisation of its protein substrates

[1,2]Carsten Jers, [2]Malene Mejer Pedersen, [2]Dafni Katerina Paspaliari, [4]Wolfgang Schütz, [2]Christina Johnsson, [2,3]Boumediene Soufi, [4]Boris Macek, [2]Peter Ruhdal Jensen and [1,2]Ivan Mijakovic*

[1] Microbiologie et Génétique Moléculaire, AgroParisTech-INRA-CNRS, FR-78850 Thiverval-Grignon, France

[2] Center for Systems Microbiology, Technical University of Denmark, DK-2800 Lyngby, Denmark

[3] Center for Biological Sequence Analysis, Technical University of Denmark, DK-2800 Lyngby, Denmark

[4] Proteome Center Tübingen, University of Tübingen, D-72076 Tübingen, Germany

*For correspondence. E-mail: Ivan.Mijakovic@grignon.inra.fr, Tel: +33(0)1 30 81 45 40, Fax: +33(0)1 30 81 54 57

This supplement contains:

Table 2.S1 and 2.S2

Figures 2.S1 to 2.S8

Table 2.S1: Oligos used in this study with underlined restriction sites.

Name	sequence	description
Primers for protein production constructs		
Ldh fwd	CGCGGATCCATGATGAACAAACATGTAAATAAAG	*Bam*HI
Ldh rev	TCCCCCCGGGTTAGTTGACTTTTTGTTCTGCAAAATG	*Cfr*9I
Eno fwd	CGCGGATCCATGCCATACATTGTTGATGTTTATG	*Bam*HI
Eno rev	TCCCCCCGGGTTACTTGTTTAAGTTGTAGAAAGAG	*Cfr*9I
oppA fwd	CGCGGATCCATGAAAAAACGTTGGTCGATTGTCAC	*Bam*HI
oppA rev	TCCCCCCGGGTTATTTAAAATATGCGTTTCTGAAATAAAC	*Cfr*9I
yvyG fwd	CGCGGATCCATGTCAGCGAAGGCAATTATTGAAC	*Bam*HI
yvyG rev	TCCCCCCGGGCTAAGCTTTTGAATCAAACAGTTTC	*Cfr*9I
Asd fwd	CGCGGATCCATGGGAAGAGGTTTACACGTAGCTG	*Bam*HI
Asd rev	TCCCCCCGGGTTATACGAGGTTTAGTTTTTTCAG	*Cfr*9I
yjoA fwd	CGCGGATCCATGTGCCAATCCAATCAAATTGTC	*Bam*HI
yjoA rev	TCCCCCCGGGCTACATGCGCTGCTGATAGAAAGG	*Cfr*9I
ynfE fwd	CGCGGATCCATGGATGAAATACTGAAACAG	*Bam*HI
ynfE rev	TCCCCCCGGGTCAAAAAAATGGTGTTGTCAATCC	*Cfr*9I
infA fwd	CGCGGATCCATGGCGAAAGACGATGTAATTGAAG	*Bam*HI
infA rev	TCCCCCCGGGTTATTTGTAACGGTACGTAATCCTG	*Cfr*9I
yorK fwd	CGCGGATCCATGGAGTATAGACTAATTGGCGACAATG	*Bam*HI
yorK rev	TCCCCCCGGGTTAAAACACAAGTTCTTTTTGTTTTG	*Cfr*9I
yclM fwd	CGCGGATCCATGAAGGTCGTTAAATTCGGAGGC	*Bam*HI
yclM rev	TCCCCCCGGGAGAGATCAGCACGCCCGCGAAAAATTC	*Cfr*9I

Construction of chimeric kinase

tkmA+ (1)	CGCGGATCCATGGATAATACGATCAAATCGGAAGAGC	*Bam*HI
ptkA+ plus *tkmA*- (2)	AGCCTCTGTTTTTTCTAAGCGCCATAGCGCCAAAATGTCCACTCCCCGTT	Fusion of TkmA and PtkA
ptkA- (3)	AAAACTGCAGTTATTTTTGCATGAAATTGTCCTTG	*Pst*I

Construction of pQE30-strep

pQE-strep+	AATTCATTAAAGAGGAGAAATTAACTATGTGGAGCCACCCGCAGTTCGAAAAAG	Linker
pQE-strep-	GATCCTTTTTCGAACTGCGGGTGGCTCCACATAGTTAATTTCTCCTCTTTAATG	Linker

Expression of fluorescent gene fusions

yfpC fwd	CGCGGATCCGGGCCCCCTAGGATGCTGAGCAAGGGCGAGGAGCTG	*Bam*HI – *Apa*I – *Avr*II
yfpC rev	TCCCCCGGGGGTACCTTACTTGTACAGCTCGTCCATGCCG	*Cfr*9I – *Kpn*I (with stop)
cfpN fwd	CGGGGTACCAAAGGAGGAAAACATTTGGATTCAATAGAAAAGGTAAGCGAATTTGCCACCATGCGTAGCAAAGGAGAAGAAC	*Kpn*I – RBS – 8 codons *comGA*
cfpN rev	CCGGAATTCCCCGGGTTTGTAGAGCTCATCCATGCCATG	*Eco*RI – *Cfr*9I (no stop)
ptkA fwd	TCCCCCGGGATGGCGCTTAGAAAAAACAGAG	*Cfr*9I
ptkA rev	CCGGAATTCTTATTTTTGCATGAAATTGTCCTTG	*Eco*RI
tkmA C fwd	CGCGGATCCAAAGGAGGAAAACATATGGGAGAATCTACAAGCTTAAAAG	*Bam*HI – RBS
tkmA C rev	GCGCTCCTAGGAGCGCCAAAATGTCCACTCCCCG	*Avr*II (no stop)
asd C fwd	CGCGGATCCAAAGGAGGAAAACATATGGGAAGAGGTTTACACGTAGCTG	*Bam*HI – RBS
asd C rev	GCGCTCCTAGGTACGAGGTTTAGTTTTTTCAGGC	*Avr*II (no stop)
infA C fwd	CGCGGATCCAAAGGAGGAAAACATATGGCGAAAGACGATGTAATTGAAG	*Bam*HI – RBS
infA C rev	GCGCTCCTAGGTTTGTAACGGTACGTAATCCTGCC	*Avr*II (no stop)
ldh C fwd	CGCGGATCCAAAGGAGGAAAACATATGATGAACAAACATGTAAATAAAG	*Bam*HI – RBS
ldh C rev	GCGCTCCTAGGGTTGACTTTTTGTTCTGCAAAATG	*Avr*II (no stop)
yjoA C fwd	CGCGGATCCAAAGGAGGAAAACATATGTGCCAATCCAATCAAATTGTC	*Bam*HI – RBS
yjoA C rev	GCGCTCCTAGGCATGCGCTGCTGATAGAAAGGAAGC	*Avr*II (no stop)

ynfE C fwd	CGCGGATCCAAAGGAGGAAAACATATGGATGAAATACTGAAACAG	*Bam*HI – RBS
ynfE C rev	GCGCTCCTAGGAAAAAATGGTGTTGTCAATCC	*Avr*II (no stop)
yorK C fwd	CGCGGATCCAAAGGAGGAAAACATATGGAGTATAGACTAATTGGCGAC	*Bam*HI – RBS
yorK C rev	GCGCTCCTAGGAAACACAAGTTCTTTTTGTTTTGTG	*Avr*II (no stop)
yvyG C fwd	CGCGGATCCAAAGGAGGAAAACATATGTCAGCGAAGGCAATTATTGAAC	*Bam*HI – RBS
yvyG C rev	GCGCTCCTAGGAGCTTTTGAATCAAACAGTTTC	*Avr*II (no stop)
oppA C fwd	CGCGGATCCAAAGGAGGAAAACATATGAAAAAACGTTGGTCGATTGTC	*Bam*HI – RBS
oppA C rev	GCGCTCCTAGGTTTAAAATATGCGTTTCTGAAATAAAC	*Avr*II (no stop)
eno-C fwd	CGCGGATCCAAAGGAGGAAAACATATGCCATACATTGTTGATGTTTATG	*Bam*HI – RBS
eno-C rev	GCGCTCCTAGGCTTGTTTAAGTTGTAGAAAGAGTTG	*Avr*II (no stop)
ssbA C fwd	CGCGGATCCAAAGGAGGAAAACATATGCTTAACCGAGTTGTATTAGTCG	*Bam*HI – RBS
SsbA C rev	GCGCTCCTAGGGAATGGAAGATCATCATCCGAGATG	*Avr*II (no stop)
ugd C fwd	CGCGGATCCAAAGGAGGAAAACATATGAATATAACAGTCATCGGAACAG	*Bam*HI – RBS
Ugd C rev	GCGCTCCTAGGTTGAATTGCACCTGACGGAACAACC	*Avr*II (no stop)

Construction of Δ*ptkA* strain

ΔptkA-up fwd	CCGCTCGAGCATCCATTATGAATGTTGAC	*Xho*I
ΔptkA-up rev	CGCGGATCCGGCGAATACGACAGCCAGGTTG	*Bam*HI
ΔptkA-down fwd	CGCGGATCCGCACAGATTTTAGGCAACGTGGC	*Bam*HI
ΔptkA-down rev	CTAGTCTAGACTGATTTCCTGTCCCGGAAGC	*Xba*I

Construction of Δ*oppA* strain

ΔoppA fwd	CCGGAATTCATGCCGATCAATAAGAAAATTGCAG	*Eco*RI
ΔoppA rev	CGCGGATCCATCCTTCTTTGTTATCCTCAAATCC	*Bam*HI

Table 2.S2: Overview of experimental data for proteins not activated by PtkA-dependent phosphorylation. Protein activities are given in % of wild type.

Protein / Condition		Untreated [%]	Phosphorylated [%]
Ldh	NADH[1,2]	100 ± 3	90 ± 4
	NADH + FBP[2]	100 ± 1	103 ± 3
	NADPH[1,3]	100 ± 0	101 ± 11
	NADPH + FBP[3]	100 ± 7	112 ± 0
Enolase		100 ± 3	109 ± 3
OppA	Exp. phase	100 ± 1	99 ± 1
	Trans. phase	100 ± 1	100 ± 2
	Δ*oppA*	0	
YvyG		100	120

[1] ~50 fold higher activity was observed with NADH compared to NADPH

[2] ~10-fold activation by FBP observed

[3] ~5-fold activation by FBP observed

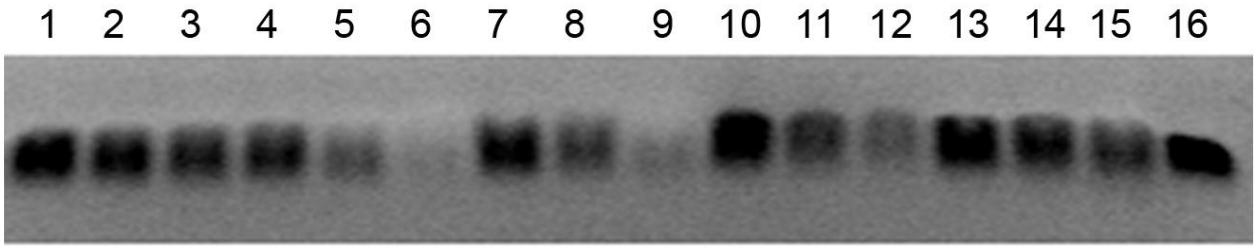


Figure 2.S1: YorK controls. YorK was incubated with the 82-mer oligonucleotide for 8, 24 and 48 h in the following contexts: alone (lanes 1-3), with TkmA and PtkA (lanes 4-6), with BSA (Bovine Serum Albumine) (7-9) and the final sample contained PtkA and TkmA, without YorK (10-12). Control with no proteins added and incubated for 48 h is in lane 13.

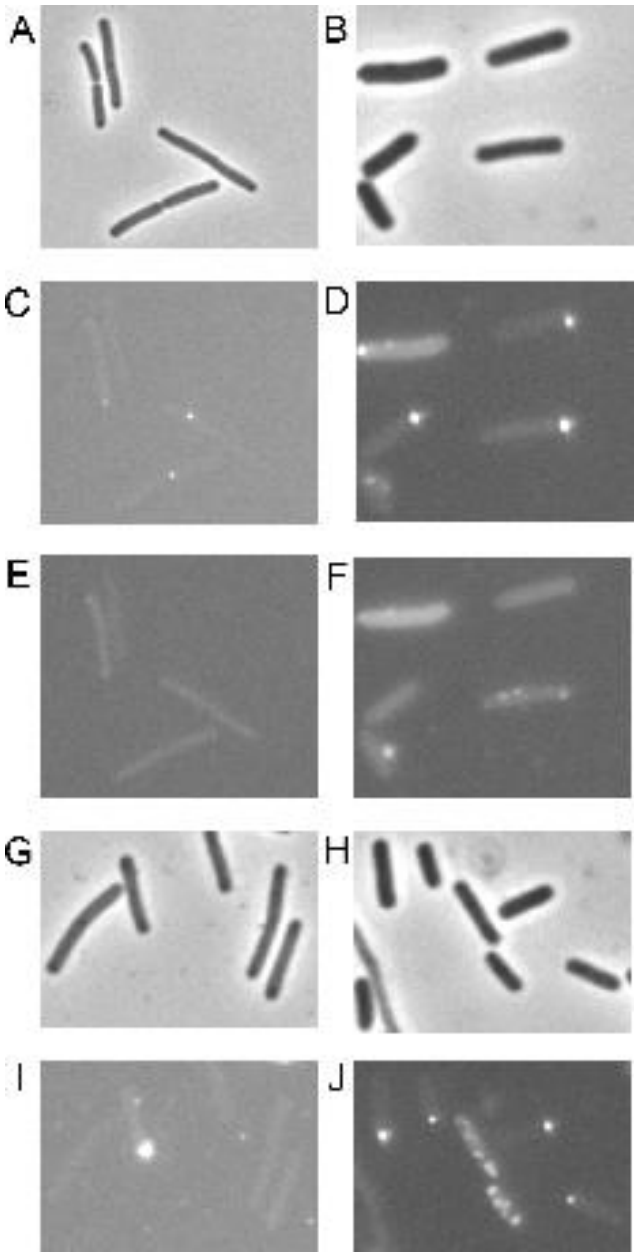


Figure 2.S2: Localisation of CFP-PtkA and Ugd-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). Ugd-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). Ugd-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

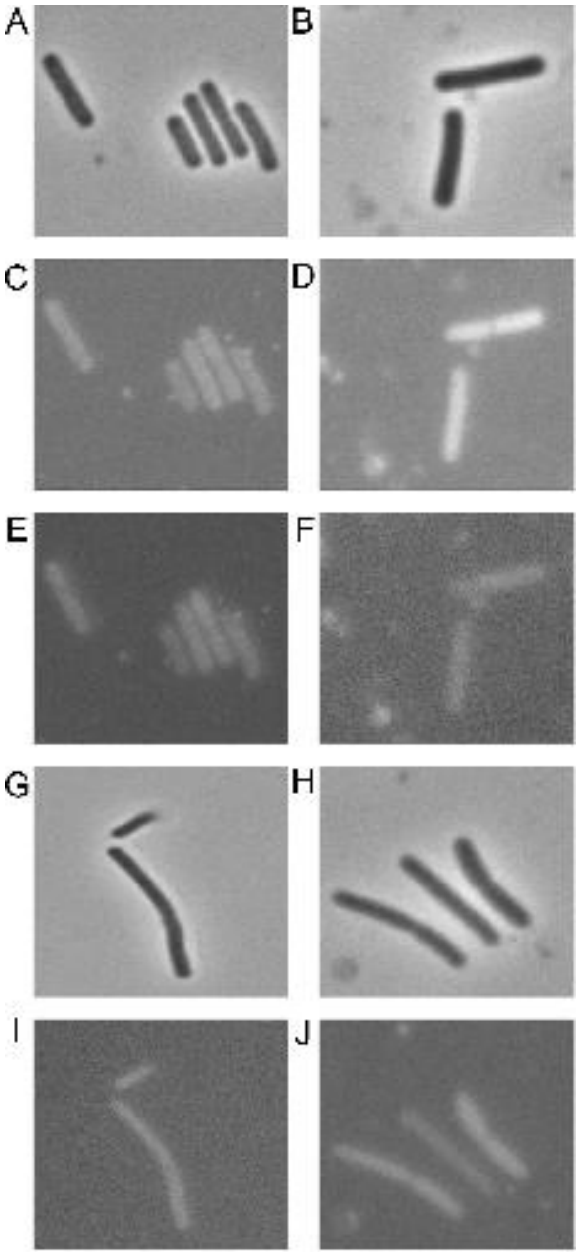


Figure 2.S3: Localisation of CFP-PtkA and Asd-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). Asd-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). Asd-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

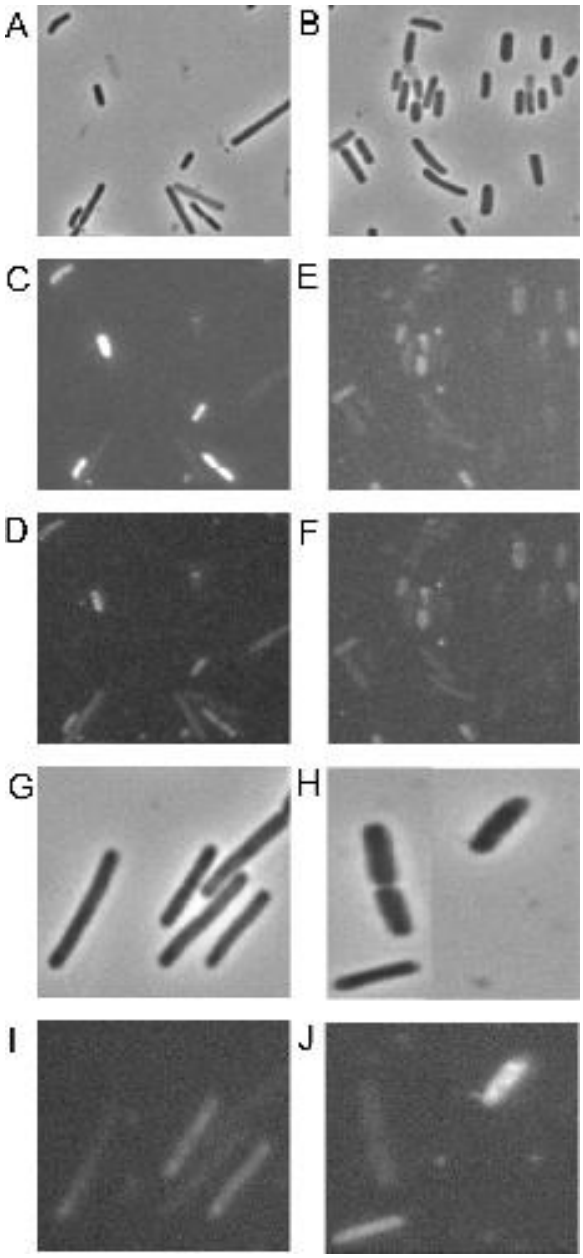


Figure 2.S4: Localisation of CFP-PtkA and YorK-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). YorK-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). YorK-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

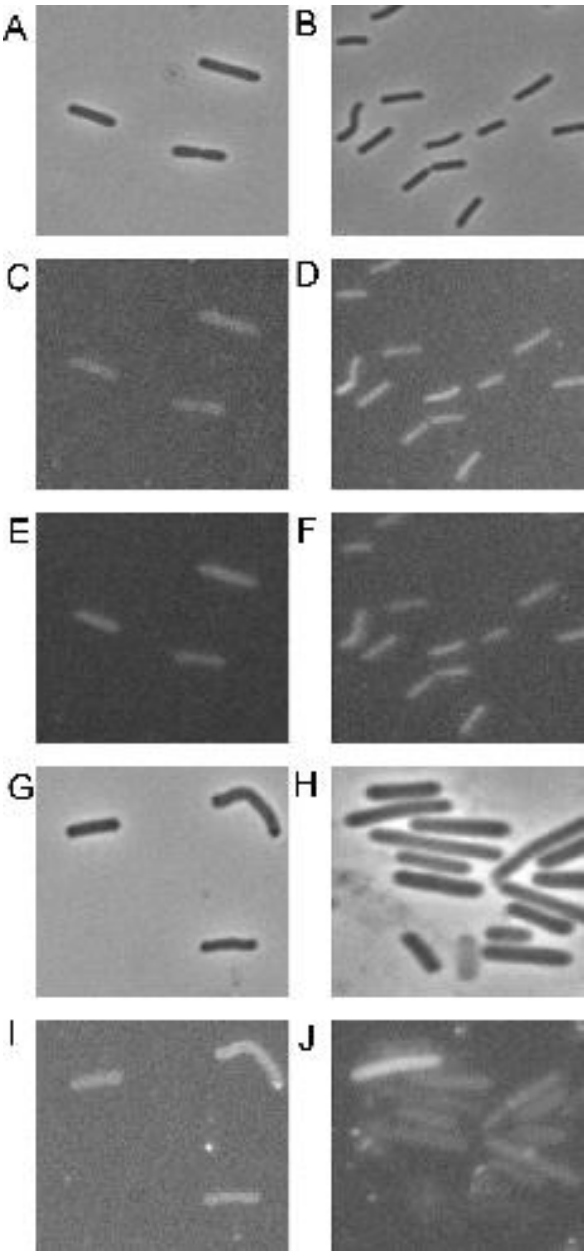


Figure 2.S5: Localisation of CFP-PtkA and InfA-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). InfA-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). InfA-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

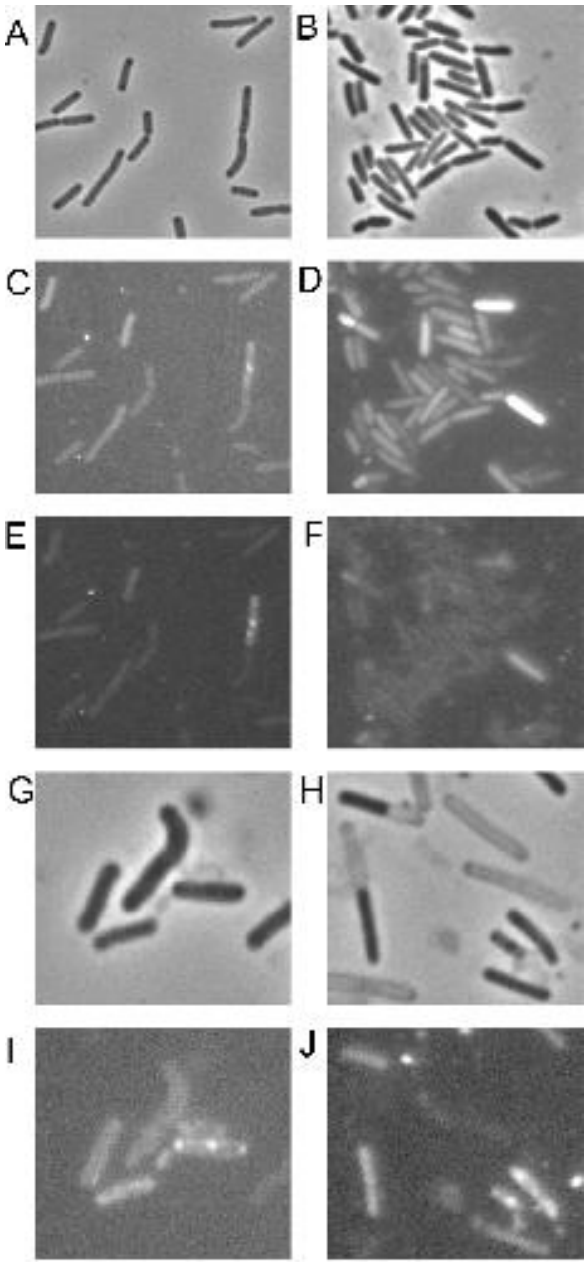


Figure 2.S6: Localisation of CFP-PtkA and YjoA-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). YjoA-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). YjoA-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

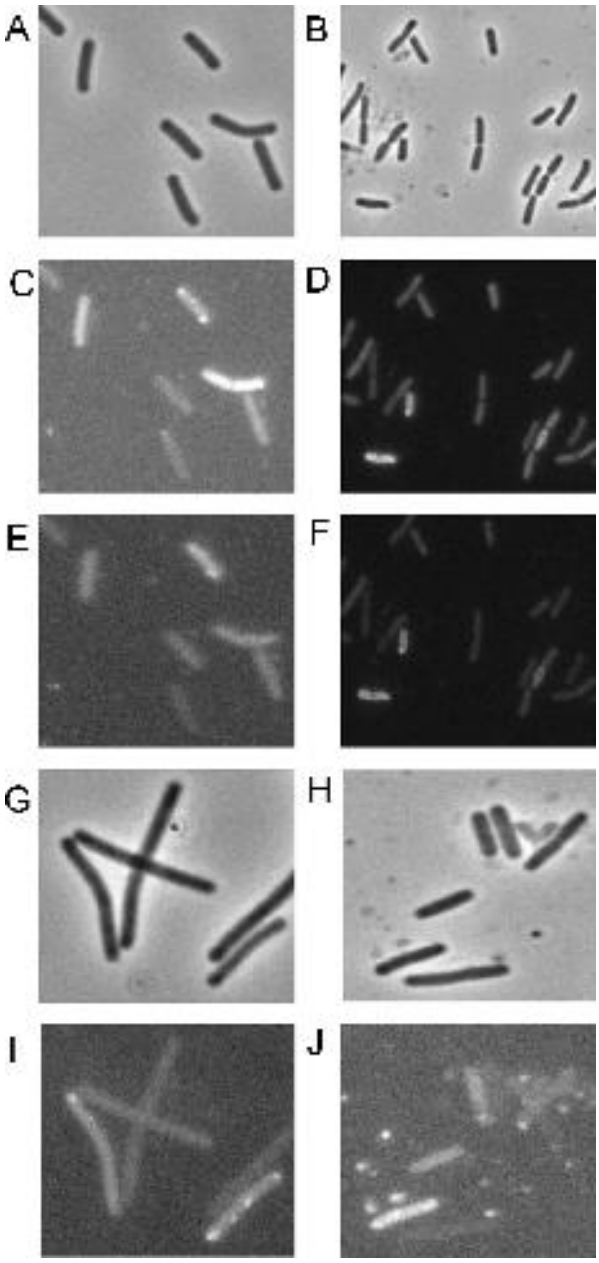


Figure 2.S7: Localisation of CFP-PtkA and Ldh-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). Ldh-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). Ldh-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

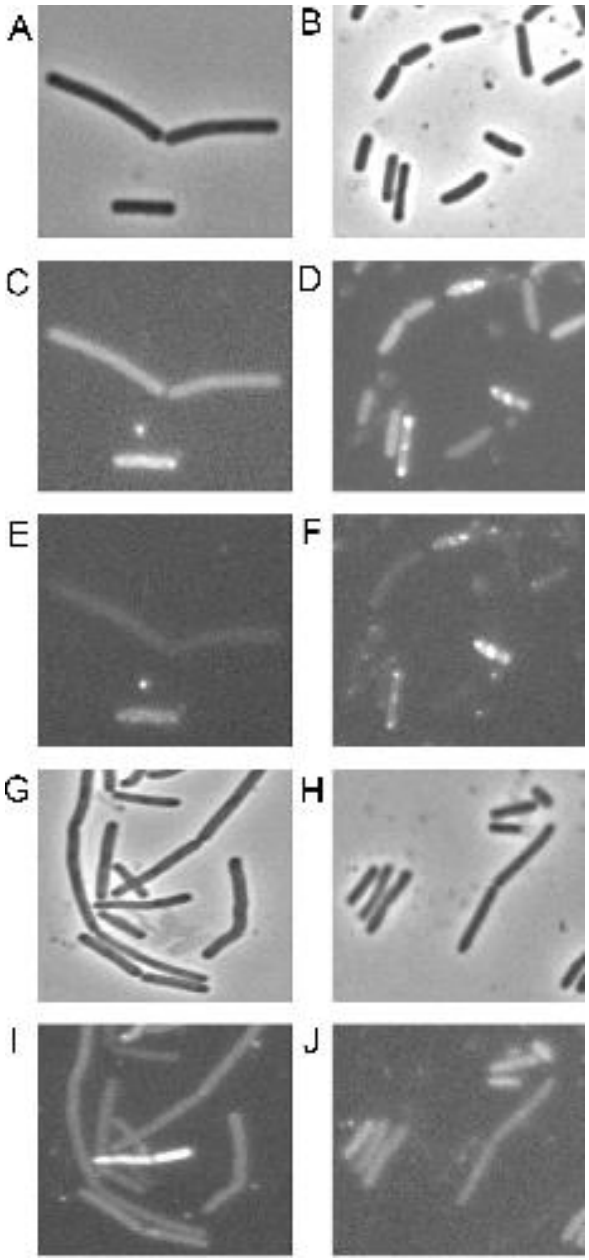


Figure 2.S8: Localisation of CFP-PtkA and YnfE-YFP. Wild type cells visualised by phase contrast in exponential (A) and stationary phase (B). YnfE-YFP visualised in exponential (C) and stationary phase (D). CFP-PtkA visualised in exponential (E) and stationary phase (F). Δ*ptkA* cells visualised by phase contrast in exponential (G) and stationary phase (H). YnfE-YFP visualised in Δ*ptkA* cells in exponential (I) and stationary phase (J).

Chapter 3

Serine phosphorylation of two-component system kinase DegS

Bacillus subtilis two-component system sensory kinase DegS is regulated by serine phosphorylation in its input domain

Carsten Jers[1], Ahasanul Kobir[2], Elsebeth Oline Søndergaard[1], Peter Ruhdal Jensen[1] and Ivan Mijakovic[2,*]

[1] Center for Systems Microbiology, Technical University of Denmark, 2800 Lyngby, Denmark

[2] Micalis UMR 1319, AgroParisTech/INRA, Domaine de Vilvert, 78352 Jouy en Josas, France

*For correspondence. E-mail: Ivan.Mijakovic@grignon.inra.fr, Tel: +33(0)1 30 81 45 40,
Fax: +33(0)1 30 81 54 57

Submitted to Molecular Microbiology

Summary

Bacillus subtilis two-component system DegS/U is notorious for the complexity of its regulation. The cytosolic sensory kinase DegS does not receive a single predominant input signal like most two-component kinases, instead it integrates a wide array of metabolic inputs that modulate its activity. The phosphorylation state of the response regulator DegU also does not confer a straightforward "on/off" response; it is fine-tuned and at different levels triggers different sub-regulons. Here we describe a novel input for the DegS kinase: serine phosphorylation of its sensing domain. We demonstrate that DegS phosphorylation can be carried out by at least two *B. subtilis* Hanks-type kinases *in vitro*, and this stimulates the phosphate transfer towards DegU. The consequences of this process were studied *in vivo*, using phosphomimetic (Ser76Asp) and non-phosphorylatable (Ser76Ala) mutants of DegS. In a number of physiological assays focused on different processes regulated by DegU, DegS S76D phosphomimetic mutant behaved like a strain with intermediate levels of DegU phosphorylation, whereas DegS S76A behaved like a strain with very low levels of DegU phosphorylation. This suggests a link between DegS phosphorylation at serine 76 and the level of DegU phosphorylation, establishing this post-translational modification as an additional trigger for this two-component system.

Introduction

Two-component systems are a ubiquitous means of signal transduction in bacteria (Hoch, 2000). The first, signal-receiving component is a sensory histidine kinase that is triggered by a stimulus binding or otherwise affecting its sensing domain. Upon activation, the histidine kinase autophosphorylates on a histidine residue, and thereafter transfers the phosphate to a specific aspartate residue of its cognate response regulator. Phosphorylation of the response regulator, in turn, triggers its regulatory role, which is in most cases transcriptional regulation via binding of a specific DNA sequence. The histidine kinases of the two-component systems are notorious for a high level of specificity, i.e. low level of cross-talk with non-cognate response regulators (Laub *et al*., 2007). Another major group of bacterial kinases involved in signal transduction are the Hanks type serine/threonine kinases (Madec *et al.*, 2003; O'Hare *et al*., 2008). Hanks type kinases and two-component histidine kinases are sometimes found fusioned in a single polypeptide in *Cyanobacteria* (Phalip *et al*., 2001), however, very few cases of crosstalk between these two protein families have been reported so far. Two recent studies pointed out that serine/threonine kinases can phosphorylate two-component response regulators: StkP from *Streptococcus pneumoniae* phosphorylates the orphan response regulator RitR (Ulijasz *et al*., 2009) and serine-threonine kinase Stk1 phosphorylates and thereby abolishes the activity of the response regulator CovR in Group B *Streptococcus* (Lin *et al*., 2009b). Interestingly, a recent phosphoproteomic study in *Bacillus subtilis*, identified the two-component system histidine kinase DegS as being phosphorylated on the residue serine76, located in the signal sensing domain (Macek *et al*., 2007). This implied the existence of a new type of crosstalk between two phosphorylation systems, namely one in which a presently unknown serine kinase would phosphorylate the two-component sensory histidine kinase.

Early mutational studies of the DegS/U two component system established that the response regulator binds DNA sequences and regulates expression of specific genes both in its phosphorylated and unphosphorylated state. This was exemplified by reciprocal effects of the two forms of DegU on exoprotease production and competence (Msadek *et al*., 1990). The importance of DegS/U was further underscored in two microarray experiments that independently demonstrated a total of 135 transcriptional units regulated either directly or indirectly by this two-component system (Ogura *et al*., 2001; Mäder *et al*., 2002). Only 22 transcriptional units were identified in the overlap between the two studies, which could indicate an even larger regulon. Whereas the initial studies of DegS/U in the laboratory strain 168 mainly focused on competence and exoprotease

production, more recent studies using an undomesticated *B. subtilis* strain demonstrated that DegS/U also affects motility, complex colony and biofilm formation. The regulation was shown to depend on the discrete levels of DegU phosphorylation, in a manner far more subtle than a simple on/off switch (Verhamme *et al*., 2007; Kobayashi, 2007). It is now well established that DegS/U plays an important role in the transition growth phase where many inputs leads to regulation on both transcriptional and post-transcriptional level. The *degSU* genes are transcribed as an operon and *degU* is itself transcribed from two additional promoters: one activated by DegU~P and the other by nitrogen starvation (Msadek *et al*., 1990; Yasumura *et al*., 2008). The signal-sensing domain of DegS interacts with the SMC-ScpA-ScpB protein complex, involved in chromosome segregation, which inhibits the kinase activity of DegS (Dervyn *et al*., 2004). Similarly, the DNA-binding activity of DegU is inhibited by RapG and this inhibition is counteracted by PhrG in response to increased cell density (Ogura *et al*., 2003). DegS/U activity is further modulated by two small regulatory peptides, DegQ and DegR. DegQ enhances the phosphotransfer from DegS~P to DegU but does not protect DegU~P from dephosphorylation (Kobayashi, 2007). The latter is accomplished by DegR that stabilises DegU~P (Mukai *et al*., 1992). DegQ is synthesized in response to quorum sensing via ComPA and has been hypothesised to be a determinant for the transition from motile to sessile-growth state (Kobayashi, 2007). *degR* expression is SigD dependent and peaks in late exponential phase, but the physiological role remains elusive (Ogura *et al*., 1997).

Despite the fact that the DegS/U two-component system is submitted to an elaborate control at both transcriptional and protein level, no specific DegS-activating signal has so far been proposed. Most two-component system histidine kinases are transmembrane proteins, presumably activated by extracellular signals via their N-terminal signal-sensing domains protruding from the cell surface. By contrast, DegS is a cytosolic protein and hence responds to, and integrates several cytosolic signals, some of which have been listed above. In this study, we examined the possibility that phosphorylation of DegS on serine 76 residue could represent a novel input for this regulatory system. We demonstrated that a phospho-mimetic mutant of DegS exhibits increased histidine-autophosphorylation *in vitro*, leading to an increased DegU~P pool, and influencing the transcription of key DegU-dependent promoters *in vivo*. We also identified two Hanks-type protein kinases capable of phosphorylating DegS *in vitro*.

Results and discussion

DegS is phosphorylated *in vitro* by *B. subtilis* Hanks-type serine/threonine kinases

In order to characterize the serine 76 phosphorylation of DegS, we first asked the question whether this phosphorylation is auto-catalyzed or it requires another kinase. To answer this question we carried out *in vitro* phosphorylation experiments with purified DegS and ^{32}P-γ-ATP. Phospho-histidine and phospho-serine can be easily distinguished, the latter being stable in acidic conditions. Since the entire radioactive label present on autophosphorylated DegS was removed by acid treatment (figure 3.1), we concluded that DegS was incapable of autophosphorylating on serine. There are a number of poorly characterized serine/threonine kinases in *B. subtilis*, mainly belonging to the family of Hanks-type kinases (Leonard *et al*., 1998). The most extensively characterized of those is the kinase PrkC that has recently been shown to participate in signalling underlying spore germination (Shah *et al*., 2008). We purified the three Hanks-type kinases PrkC, YabT and YbdM, and a putative kinase PrkA to test their ability to phosphorylate DegS *in vitro*. PrkA (data not shown) and PrkC were unable to phosphorylate DegS, whereas both YbdM and YabT tested positive for DegS phosphorylation (figure 3.1). In order to verify whether serine 76 is indeed the residue phosphorylated by these kinases, we constructed a mutant protein with a non-phosphorylatable replacement DegS S76A. Phosphorylation of DegS S76A by YbdM was completely abolished, suggesting that it is the major phosphorylation site, and that the kinase YbdM is specific for this site (figure 3.1A). In the same assay, phosphorylation of DegS S76A by YabT was as efficient as that of the wild type (figure 3.1B). This situation is not unprecedented (Petranovic *et al*., 2009), and residual phosphorylation in such case could be due to a presently unknown secondary site, or the lack of specificity exhibited by the kinase under *in vitro* conditions. Our conclusion was that *in vitro* at least two different Hanks-type kinases can phosphorylate DegS, of which YbdM is specific for the residue serine 76. DegS is also known to exhibit phosphatase activity, so we tested whether serine phosphorylation of DegS could be removed by the protein itself. For this, we removed the ATP from the phosphorylation reactions catalyzed by YbdM and YabT, and allowed the dephosphorylation reaction to occur for 2 h. The radiolabel on DegS remained stable, indicating the absence of phospho-serine phosphatase activity (figure 3.1A and 3.1B, lanes 9). Regarding the situation *in vivo*, the kinase-encoding genes *yabT*, *ybdM* and *prkA* have been reported to belong to the sigma F, G and E regulon respectively (Wang *et al*., 2006; Steil *et al*., 2005) hence linking them to sporulation specific processes. However, a recent

transcriptomics study (Rasmussen *et al*., 2009) identified *yabT*, *ybdM* and *prkC* as transcribed in the exponential growth phase, which makes it highly probable that these kinases are present in the transition phase when DegS activity is triggered.

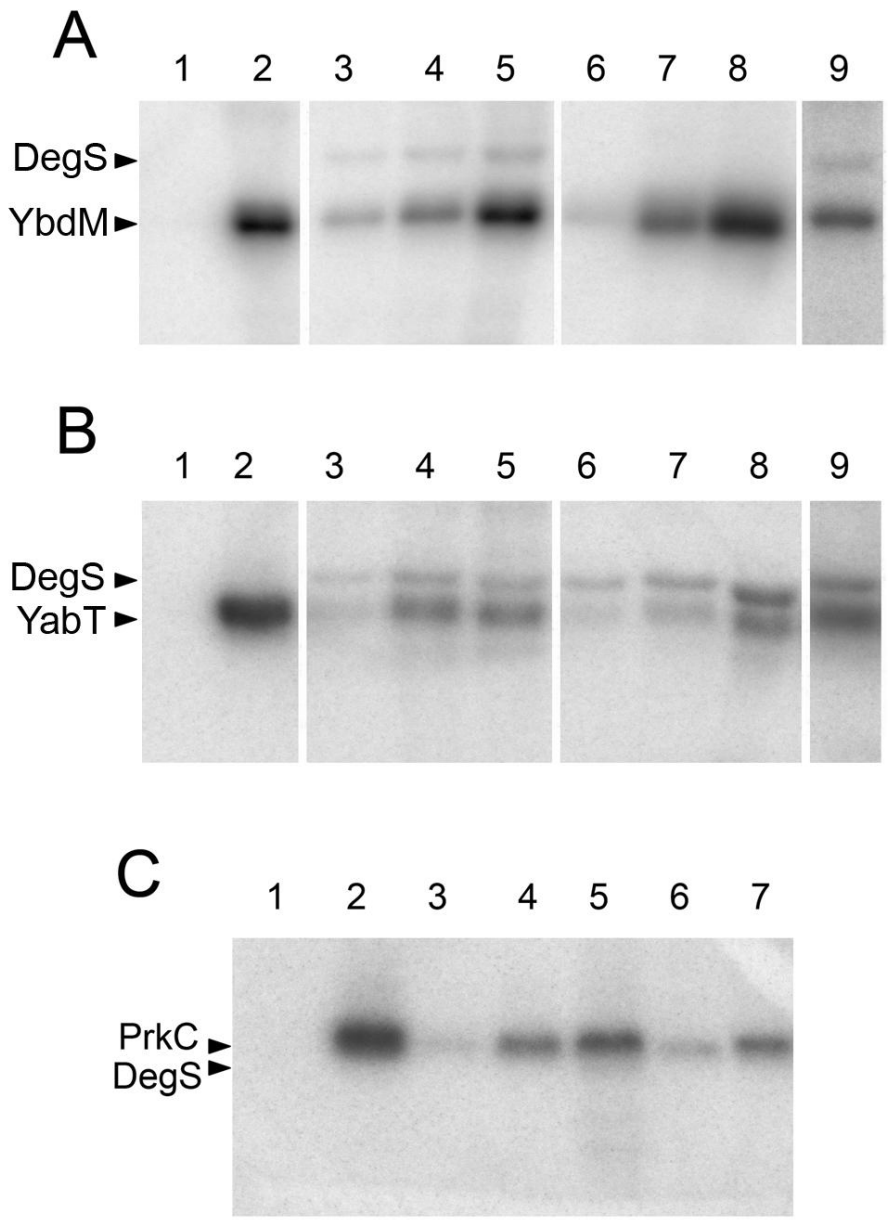


Figure 3.1: Phosphorylation assays *in vitro*: autoradiography of SDS-Polyacrylamide gels with proteins that had been incubated with ^{32}P-γ-ATP. (A+B) Phosphorylation of DegS by the kinase YbdM (A) or YabT (B). Lanes 1 and 2 are controls, with DegS alone and kinase alone, respectively. Lanes 3-5 show phosphorylation of DegS by kinase for 15, 30 and 60 min, respectively. Lanes 6-8 show phosphorylation of DegS S76A by kinase for 15, 30 and 60 min, respectively. Lane 9 shows the equivalent of the reaction from lane 5, after desalting to remove the ATP and a 2 h incubation to test for phosphatase activity. (B) (C) Phosphorylation of DegS by the kinase PrkC. Lanes 1 and 2 are controls, with DegS alone and PrkC alone, respectively. Lanes 3-5 contain the reactions where PrkC concentration has been varied (2, 4 and 10 nM, respectively), and in lanes 6-7 the pH has been varied (pH 7 and pH 8, respectively).

Phosphomimetic mutant DegS S76D exhibits increased autophosphorylation and phosphotransfer to DegU *in vitro*

Rarely more than several percent of the target protein is phosphorylated during *in vitro* kinase assays, unless a specific kinase-activating signal is present (Mijakovic *et al*., 2003). In order to study the regulatory effects of phosphorylation, phosphomimetic mutants, with the phosphorylatable residue replaced by a larger, negatively charged amino acid are often employed (Wittekind *et al*., 1989). Since our data indicated that DegS can be phosphorylated by two Hanks-type kinases, for which the specific effectors (or conditions) that trigger their activity towards DegS are unknown, it seemed particularly promising to study the effects of DegS phosphorylation *in vivo* using a phosphomimetic mutant DegS S76D. Before using the mutant protein *in vivo*, we checked its behaviour by an *in vitro* phosphorylation assay. Purified DegS S76D showed an increase in DegU phosphorylation on aspartate (figure 3.2A) as well as an increase in autophosphorylation on histidine (figure 3.2B) compared to the wild type. The respective activities of DegS S76A were above

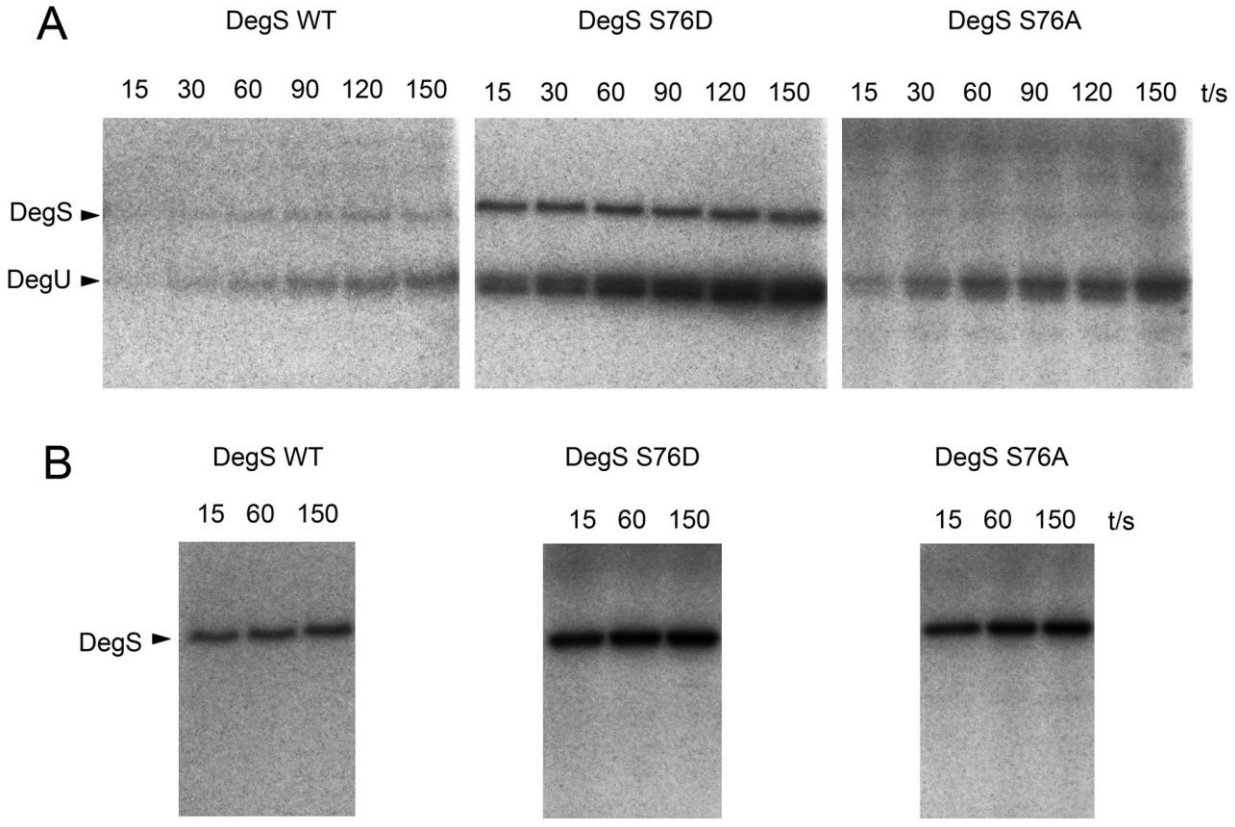


Figure 3.2: DegS autophosphorylation and DegU phosphorylation *in vitro*: autoradiography of SDS-Polyacrylamide gels with proteins that had been incubated with ^{32}P-γ-ATP. Time periods are indicated for each lane, for DegU phosphorylation (A) and DegS autophosphorylation (B).

the wild type level, but below that of DegS S76D (figure 3.2). Interestingly, DegS S76D maintained a considerable level of incorporated phosphate in the presence of DegU, whereas DegS wild type and DegS S76A incorporated much less phosphate under these conditions. DegS/U exerts a very complex regulation of several physiological processes, some of which are affected by high and others by low levels of DegU phosphorylation. We thus hypothesized that strains where wild type DegS would be replaced by DegS S76D should be unable to activate processes that are normally stimulated by non-phosphorylated DegU, and probably overly stimulate processes that require phosphorylated DegU. We therefore decided to focus on several functions that exemplify these contrasting situations to examine the regulatory role of DegS phosphorylation on serine 76.

DegS S76D negatively affects competence development

Competence development is a complex process regulated partly by DegS/U via the transcriptional activation of *comK* exerted by unphosphorylated DegU (Hamoen *et al*., 2003). To test the effect of phosphorylation of DegS serine 76 on this process, strains were constructed in which the chromosomal version of *degS* was replaced by copies encoding either DegS S76D or DegS S76A. Competence of the resulting strains was initially compared to the wild type by using a two-step transformation protocol (Yasbin *et al*., 1975). The competence of *degS* S76A (155 +/- 7 %) was about 50 % higher than in wild type (100 +/- 12 %) while the *degS* S76D mutant exhibited an approximately 5-fold reduction(18 +/- 0 %), concurring with our working hypothesis. Next, competence development on single cell level was assayed by introducing GFP under control of the *comK* promoter. In this system we further assayed the effects of inactivating individual serine/threonine kinase-encoding genes *prkC, ybdM and yabT* (figure 3.3). The overall picture was confirmed with an approximate 5-fold competence reduction in *degS* S76D, while in this set up no difference between wild type and the *degS* S76A strain was observed. A mutant of the kinase phosphorylating DegS would be expected to behave as the non-phosphorylatable *degS* S76A in this setup. All kinase mutants had competence level identical or superior to wild-type, which is in agreement with our hypothesis, but of limited evidential value, since *degS* S76A was itself non-distinguishable from the wild type. Competence in *B. subtilis* is a bistability phenomenon that occurs in a sub-population of cells which become transiently competent during a certain window of time in the transition growth phase (Maamar *et al*., 2005). We thus asked the question whether the effect of *degS* S76D on competence could be due to a temporal shift in the window of competence or an overall decrease in DegU-dependent *comK* expression. In order to determine this, *lacZ*

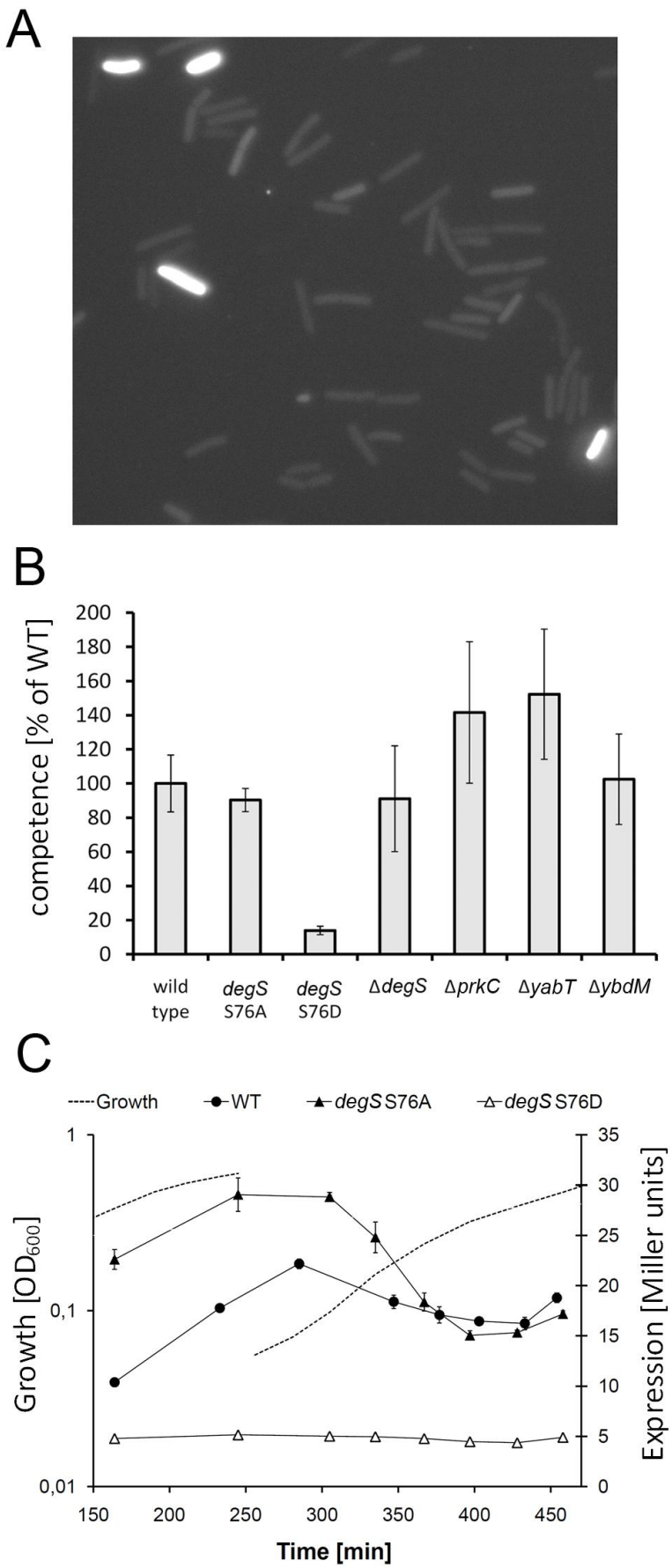


Figure 3.3: Competence assays. (A) Single cell analysis of competence: a representative picture demonstrating the difference in fluorescence intensity observed between competent and non-competent cells.(B) Single cell analysis of competence: numerical data. Competence of wild type, *degS* mutants and Ser/Thr kinase mutants normalised with respect to the wild type cells. The results (with standard deviation bars) are the average of three independent experiments. (C) *comK* promoter activity in wild type, *deg*S S76A and *degS* S76D strains grown in competence media. Wild type is shown in filled circles, *degS* S76A strain in filled triangles and *degS* S76D in open triangles. Culture growth is indicated with the dotted line (broken line indicates the dilution in the new medium). The results (with standard deviation bars) are the average from three technical replicas.

encoding β-galactosidase was placed under the control of the *comK* promoter and introduced in the *amyE locus* of wild type and mutant strains. Activity of the *comK* promoter was assayed in the two-step competence media. Expression of *comK* was similar in the wild type and *degS* S76A, peaking out as expected in the early transition phase. In the *degS* S76D mutant, expression of *comK* was 5-encoding β-galactosidase was placed under the control of the *comK* promoter and introduced in the *amyE locus* of wild type and mutant strains. Activity of the *comK* promoter was assayed in the two-fold lower, and not induced at all in the transition phase. Since *comK* expression is known to be activated by unphosphorylated DegU, the conclusion we reached based on the presented data was that the *degS* S76D mutation indeed leads to an increase in the DegU~P pool *in vivo*, in accordance to our phosphorylation data obtained *in vitro*.

DegS76D affects complex colony formation and swarming

After confirming the effect of the phospho-mimetic mutation S76D on the pool of DegU~P using the *comK* promoter, activated by unphosphorylated DegU, we set out to confirm this finding from the opposite angle. We next examined the *yvcA* promoter, recently shown to be activated by intermediate levels of DegU~P (Verhamme *et al*., 2007). We used the same experimental setup with promoter-*lacZ* fusions introduced ectopically, and the promoter activities were determined in cells grown in LB medium (figure 3.4). A significant increase in *yvcA* promoter activity was observed in the *degS* S76D strain compared to the wild type and *degS* S76A strain. A gradual increase in expression from the *yvcA* promoter in the wild type *B. subtilis* compared to the *degS* S76A strain observed in the stationary phase could possibly reflect a gradual increase in the level of DegS serine phosphorylation, which is completely abolished in *degS* S76A. These results further substantiate that the level of DegU~P is increased in the *degS* S76D mutant. The *yvcA* promoter has been reported to be inhibited at high

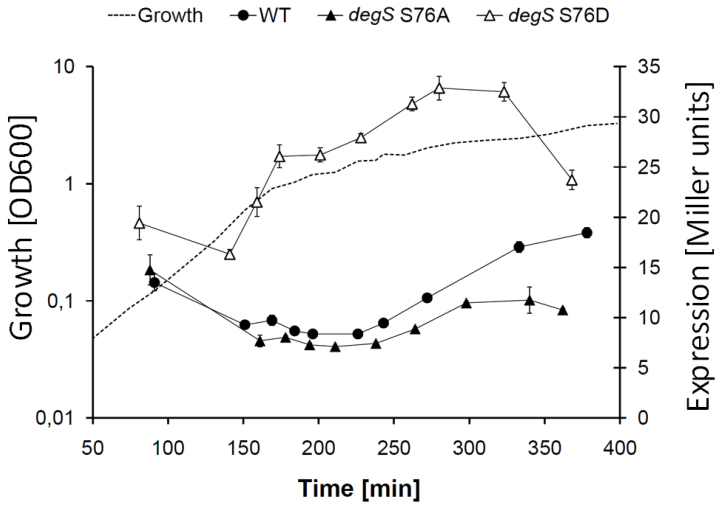
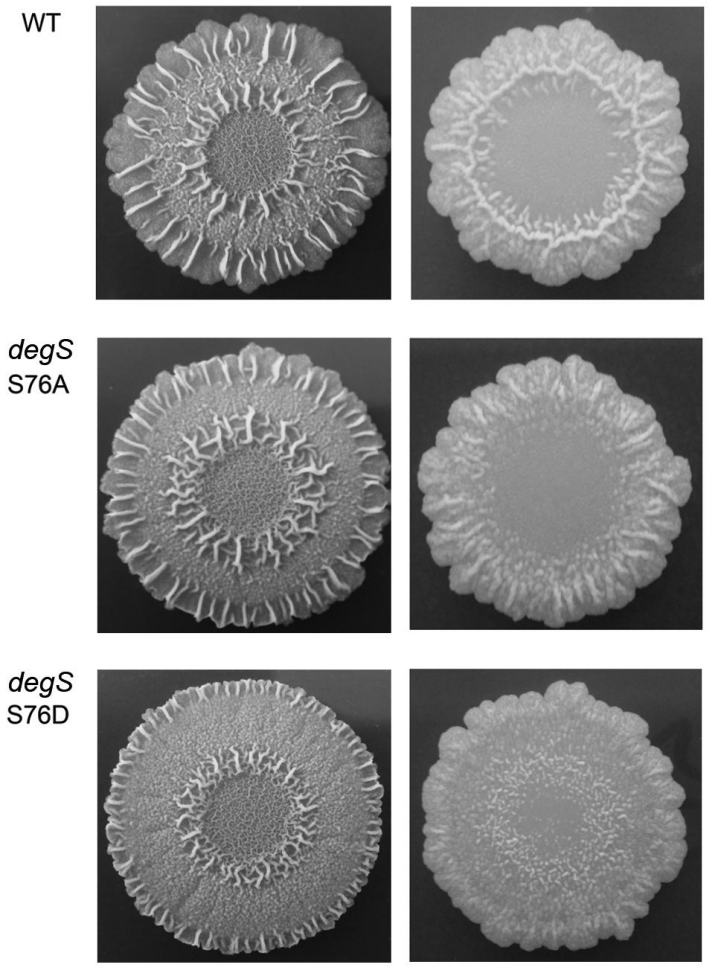


Figure 3.4: *yvcA* promoter activity in wild type, *deg*S S76A and *degS* S76D strains grown in LB medium. Wild type is shown in filled circles, *degS* S76A strain in filled triangles and *degS* S76D in open triangles. Culture growth is indicated with the dotted line. The results (with standard deviation bars) are the average from three technical replicas.

levels of DegU~P, indicating that this mutation leads rather to intermediate than excessive amount of DegU~P. Further substantiating this, protease production, known to be activated by high levels of DegU~P, was not stimulated by the *degS* S76D mutant (data not shown).

YvcA has been shown to play an important role in complex colony formation (Verhamme *et al*., 2007). The effects on *yvcA* promoter activity prompted us to investigate the effects of the *degS* S76D mutation on this and the other social behavioural traits pellicle formation and swarming that are regulated by low levels of DegU~P. The laboratory strain *B. subtilis* 168 does not readily swarm due to defects in surfactin production and a frame shift mutation in *swrA* (Kearns *et al*., 2004) and we therefore tested these traits in the undomesticated strain NCIB 3610. No effect was observed on pellicle formation (data not shown). Concerning complex colony formation, all strains exhibited some variability on the MSgg medium. However, a clear trend was that wild type and *degS S*76A colonies were capable of producing larger and more complex aerial structures than the strain *degS S*76D (figure 3.5). On LB there was less variation, and both *degS* S76A and *degS* S76D seemed to be impaired with respect to forming complex colonies compared to the wild type (figure 3.5).



	Colony diameter [cm]	
	MSgg	LB
WT	2.09	1.12
degS S76A	1.95	1.12
degS S76D	1.79	1.24

Figure 3.5: Complex colony formation of the wild type *B. subtilis* and mutant strains *degS* S76A and *degS* S76D, grown on MSgg medium (left) and LB medium (right). Representative colonies are shown, together with the colony size for reference.

A stronger phenotype was observed in swarming (figure 3.6). While wild type and *degS* S76A swarming pattern and speed were similar, the *degS* S76D strain exhibited notable differences (figure 3.6A). In the

degS S76D strain the onset of swarming was delayed for at least 60 min. Thereafter it swarmed at normal speed until reaching about one half the plate, and then it paused, followed by what appeared to be a consolidation phase. After about 2 hours it continued swarming and went on to reach the end of the plate. Due to this stop-and-go behaviour, the *degS* S76D final swarm exhibited concentric layers at various stages of consolidation (figure 3.6B). It is difficult to explain this altered swarming pattern in the strain *degS* S76D. It has previously been reported that expression of the flagella operon is inhibited at high DegU~P levels (Amati *et al*., 2004), but no difference in flagella amount or organisation in the three strains was observed at the time of *degS* S76D consolidation (data not shown). Further, the phenotype does not seem to be related to either surfactin production or the *swrA* mutation since a lab strain cured for these mutations, DS155 (Kearns *et al*., 2004), exhibits no phenotype on swarming (data not shown). Nevertheless, these data collectively support the idea that *degS* S76D mutation indeed leads to an increased DegU~P level, and indicate that the serine phosphorylation event could have the potential to regulate social behaviours regulated at low DegU~P levels in *B. subtilis*.

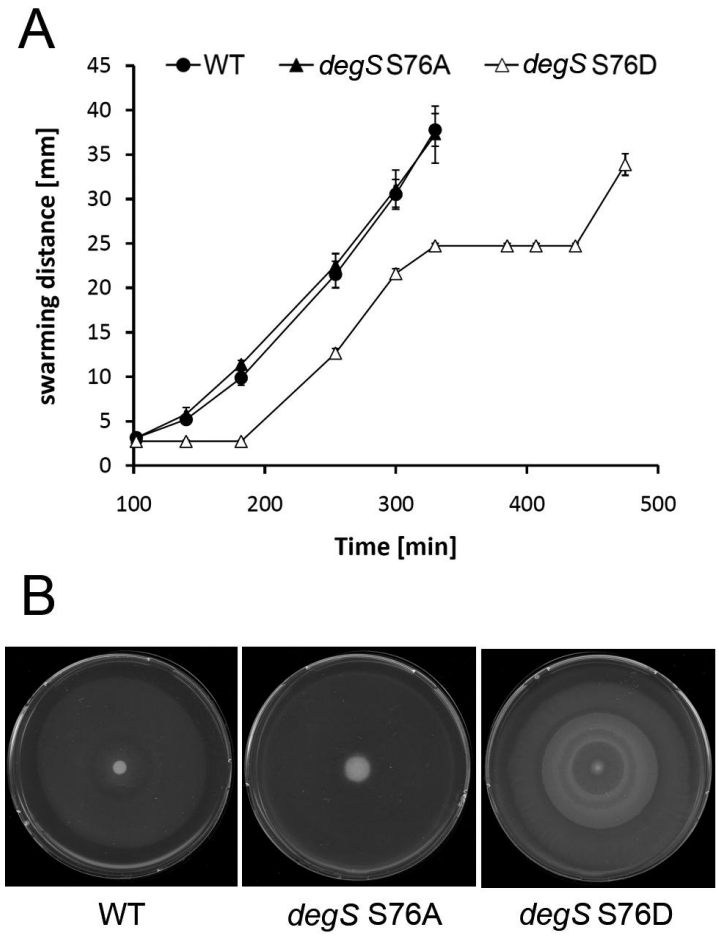


Figure 3.6: Swarming of the wild type *B. subtilis* and mutant strains *degS* S76A and *degS* S76D, Swarming speed (A) was followed by measuring the radii of the swarming zones on plates at designated time intervals. Wild type is shown in filled circles, *degS* S76A strain in filled triangles and *degS* S76D in open triangles. Consolidation of swarms (B) was documented one hour after the swarming reached the end of plates (367 min for wild type, 374 min for *degS* S76A, and 502 min for *degS* S76D). One of three independent experiments (all yielding similar results) is shown.

Concluding remarks

Here we describe, to the best of our knowledge, the first example of a bacterial two component sensory kinase that is regulated via serine phosphorylation of its input domain by a Hanks type Ser/Thr kinase. DegS was phosphorylated by Hanks type serine/threonine kinases YbdM and

YabT *in vitro* and a *degS* S76D mutation lead to higher level of DegU~P compared to the wild type *in vivo*. This suggests that phosphorylation of serine 76 of DegS contributes to an already very complex process of regulating the level of DegU phosphorylation in *B. subtilis*. It seems unlikely that all Hanks type kinases present in *B. subtilis* equally contribute to phosphorylating DegS serine 76. However, these kinases are presently almost completely uncharacterised (except PrkC) and more work will be needed to elucidate specific signals that trigger their expression, activity and substrate specificity. It is not at present clear what cellular signal might control DegS serine phosphorylation. The residue, however, is phosphorylated in late exponential growth phase (Macek *et al*., 2007). Our results indicated that DegS serine phosphorylation corresponds to an intermediary level of DegU phosphorylation *in vivo*, pointing towards a possible role in regulating phenomena such as motility and complex colony formation.

Experimental procedures

Bacterial strains and growth conditions

E. coli NM522 was used for plasmid propagation in cloning experiments. The chaperone overproducing strain *E. coli* M15 carrying pREP4-GroESL (Amrein *et al*., 1995) was used for protein synthesis. *B. subtilis* strains used in this study are listed in table 3.1. *E. coli* and *B. subtilis* cells were grown at 37 °C with shaking in LB medium. In addition, *B. subtilis* was grown in competence media for transformation experiments as described (Yasbin *et al*., 1975) and MSgg medium for complex colony experiments (Branda *et al*., 2001). For *E. coli* ampicillin (100 µg/mL), kanamycin (25 µg/mL), tetracycline (8 µg/mL) and for *B. subtilis* erythromycin (5 µg/mL), neomycin (5 µg/mL) and tetracycline (15 µg/mL) were added as appropriate.

DNA manipulations and strain construction

B. subtilis genes *degS, degU, prkA, prkC* (catalytic domain), *yabT* (catalytic domain) and *ybdM* were PCR-amplified using genomic DNA from the strain 168 as template. In order to improve solubility of PrkC and YabT only the cytosolic part containing the active site was used. Point mutations *degS* S76A and S76D were obtained using two partially overlapping mutagenic primers (table 3.2). The PCR products were inserted in the vector pQE30 (Qiagen) using appropriate

Table 3.1: List of *B. subtilis* strains used in this study

Strain	Description	Reference
168		(Kunst *et al*., 1997)
168-*degS* NS	*degS* K9stop	This work
168-PcomK	*amyE*::P*comK-lacZ*	This work
168-PcomK-gfp	*amyE*::P*comK-gfp*	This work
168-PyvcA	*amyE*::P*yvcA-lacZ*	This work
degS S76A	*degS* S76A	This work
S76A-PcomK	*degS* S76A *amyE*::P*comK-lacZ*	This work
S76A-PcomK-gfp	*degS* S76A *amyE*::P*comK-gfp*	This work
S76A-PyvcA	*degS* S76A *amyE*::P*yvcA-lacZ*	This work
degS S76D	*degS* S76D	This work
S76D-PcomK	*degS* S76D *amyE*::P*comK-lacZ*	This work
S76D-PcomK-gfp	*degS* S76D *amyE*::P*comK-gfp*	This work
S76D-PyvcA	*degS* S76D *amyE*::P*yvcA-lacZ*	This work
168-ΔdegS PcomK-gfp	Δ*degS*::pMUTIN2 *amyE*::P*comK-gfp*	This work
168-ΔprkC PcomK-gfp	Δ*prkC*::pMUTIN2 *amyE*::P*comK-gfp*	This work
168-ΔybdM PcomK-gfp	Δ*ybdM*::pMUTIN2 *amyE*::P*comK-gfp*	This work
168-ΔyabT PcomK-gfp	Δ*yabT*::pMUTIN2 *amy*E::P*comK-gfp*	This work
3610	*sfp+ swrA+*	Bacillus Genetic Stock Center
3610 degS S76A	*sfp+ swrA+ degS* S76D	This work
3610 degS S76D	*sfp+ swrA+ degS* S76D	This work
DS155	PY79 *sfp+ swrA+*	Kearns *et al*., 2004

restriction sites. For promoter-*lacZ* fusions the promoter regions were PCR-amplified from genomic DNA and inserted between the *Eco*RI and *Bam*HI sites of pDG268-neo (Christiansen *et al*., 1997). For *comK* promoter-*gfp* fusion, pDG268neo-PcomK was restricted with *Bam*HI and *Pci*I to remove *lacZ*, *gfp* was PCR-amplified using plasmid pFH2191 (Østergaard *et al*., 2001) as template, restricted and ligated with the vector. *B. subtilis* was transformed with the constructs yielding strains with promoter-*lacZ* or *-gfp* fusions inserted in the *amyE* locus using a one-step transformation method (Jarmer *et al*., 2002). A nonsense mutation of *degS* (K9stop) was constructed using two partially overlapping mutagenic primers. The PCR product and pHT315 (Arantes and Lereclus, 1991) were restricted with *Eco*RV and *Pvu*II and the fragments ligated. The resulting vector was devoid of the Gram-positive origin of replication. Upon transformation and integration on the chromosome, the mutation was verified by sequencing. Inactivation of *prkA*, *prkC*, *ybdM* and *yabT* was done using pMUTIN2 (Vagner *et al*., 1998). The vector pG$^+$host8

Table 3.2: Primers used in this study. Restriction sites are underlined and changed codons are in bold.

Name	Sequence	Description
Gene amplification		
DegU fwd	CGCCGCGGATCCATGACTAAAGTAAACATTGTTATT	*Bam*HI
DegU rev	CGCAATGGTACCTTATCTCATTTCTACCCAGCCATTTTT	*Kpn*I
DegS fwd	CGCCGCGGATCCATGAATAAAACAAAGATGGATTCC	*Bam*HI
DegS rev	CGCAATGGTACCTTAAAGAGATAACGGAACCTTAATCAT	*Kpn*I
PrkC fwd	GAAGATCTATGCTAATCGGCAAGCGGATCAGCGGGCG	*Bgl*II
PrkCtrunc rev	AAAACTGCAGTTACAAAACCCACGGCCACTTTTTTCTTTTTGCCG	*Pst*I, amplification of aa 1-333
YabT fwd	GAAGATCTATGATGAACGACGCTTTGACGAGTTTGGC	*Bgl*II
YabTtrunc rev	AAAACTGCAGTTAGATAAGCGTTGTTTCAAATAACCCC	*Pst*I, amplification of aa 1-321
YbdM fwd	CGGGATCCATGGCATTAAAACTTCTAAAAAAACTGC	*Bam*HI
YbdM rev	AAAACTGCAGTTATGTGACCGATTGAATGGCCCG	*Pst*I
PrkA fwd	CGGGATCCATGGATATATTAAAGAAAATTGAAAAGTAC	*Bam*HI
PrkA rev	AAAACTGCAGTTATCGGTTCAGCAGGCTGCCG	*Pst*I
RBS-gfp fwd	CGGGATCCAAAGGAGGAAAACATATGTCTAAAGGTGAAGAACTG	*Bam*HI
gfp rev	CCATACATGTTTATTTATACAGCTCATGCATGC	*Pci*I
degS NS1 fwd	CGGATATCATCTCGTGTTCTCCCGCTTC	*Eco*RV, anneals 263 bp upstream of *degS*
degS NS2 rev	GTCCAGCTGTTCATACTGCTGGCGTGACTGC	*Pvu*II, anneals 123 bp inside *degS*
prkC_ MUT_fwd	CCCAAGCTTAAAGATCCTTTTCATCGCTACG	*Hind*III
prkC_MUT_rev	CGCCCGCGGGGTGACCGTGGCGCCTTCTTTGAC	*Sac*II
yabT_MUT_fwd	CCCAAGCTTATGCAATGGAATACATAAAAGGG	*Hind*III
yabT_MUT_rev	CGCCCGCGGTTGAAGCAGCGGGTTTCCTTCG	*Sac*II
ybdM_MUTfwd	CCCAAGCTTGAATTCATCATAGACGGACAGG	*Hind*III
ybdM_MUTrev	CGCCCGCGGCAGCAAGAATAACAGCGTTTCTCC	*Sac*II
Mutagenic primers for degS		
S76A fwd	AAACCGTTTA**GCC**GAGGTCAGCCGTAATTTTCA	S76A
S76A rev	GGCTGACCTC**GGC**TAAACGGTTTCTCGCATGGC	S76A
S76D fwd	AAACCGTTTA**GAC**GAGGTCAGCCGTAATTTTCA	S76D
S76D rev	GGCTGACCTC**GTC**TAAACGGTTTCTCGCATGGC	S76D
degS NS1 rev	AATCCAGCAC**TTA**GGAATCCATCTTTGTTTTATTC	K9Stop
degS NS2 fwd	GATGGATTCC**TAA**GTGCTGGATTCTATTTTGATG	K9Stop
Promoter amplification		
PcomK fwd	CGGAATTCTAAAGAATCCCCCCAATGCC	*Eco*RI
PcomK rev	CGCGGATCCGTCTGTTTTCTGACTCATAT	*Bam*HI
PyvcA fwd	CGGAATTCGAACGCCAAGCGGAAATGCC	*Eco*RI
PyvcA rev	CGCGGATCCCCTGTCAGGGCAAGTAATAAG	*Bam*HI

containing a temperature sensitive origin of replication was used to introduce *degS* S76A and S76D point mutations *in situ*, replacing the chromosomal version of *degS* in *B. subtilis* 168 (Maguin *et al*., 1996). *Bam*HI-*Cfr*9I fragments from pQE30-*degS* S76A and S76D containing the mutated gene were inserted into pG^+host8 between the *Bam*HI and *Cfr*9I sites. *B. subtilis* was transformed with the constructs, plated on tetracycline-containing plates and incubated at the non-replicative temperature 37 °C, which selects for integration of the vector on the chromosome by single crossing-over. The transformants were further cultured in liquid LB for loss of plasmid by the second crossing-over event and the chromosomal mutation was verified by sequencing. In our hands the second-crossing over happened with a low frequency indicating that *B. subtilis*, contrary to *Lactococcus lactis*, was not severely affected by a chromosomal copy of pG^+host8 actively replicating. Point mutations in *B. subtilis* NCIB3610 were introduced by same method except that transformation with pG^+Host8 was done by PEG treatment of protoplasts (Chang and Cohen, 1979).

Protein synthesis and purification

6xHis-tagged proteins were synthesised in the chaperone-overproducing strain *E. coli* M15 carrying pREP4-groESL. Cultures were grown shaking at 37 °C to OD_{600} 0.5, induced with 1 mM IPTG and grown for an additional 3 hours. Cells were disrupted by sonication and 6xHis-tagged proteins were purified on Ni-NTA columns (Qiagen) according to manufacturer's instruction, desalted with PD-10 columns (GE-Healthcare) and stored in a buffer containing 50 mM Tris-Cl pH 7.5, 100 mM NaCl and 10 % glycerol. Protein concentrations were estimated using the Bradford assay (Bio-Rad) with BSA as standard.

***In vitro* phosphorylation assay**

Phosphorylation reactions were performed in a total volume of 30 µl, essentially as described (Mijakovic *et al*., 2003), with 180 nM DegS, DegS S76A or DegS S76D and 15 µM DegU. For serine phosphorylation of DegS, reactions contained 10 nM of either PrkA, PrkC (catalytic domain), YabT (catalytic domain) or YbdM (unless otherwise specified in the figure legend). Besides the proteins, the reaction mixture contained 50 µM ^{32}P-γ-ATP (20 µCi/mmol), 42.5 mM Tris-Cl (pH 7.5), 5 mM $MgCl_2$, 85 mM NaCl and 8.5 % glycerol. For PrkC, we varied the pH value of the Tris-Cl buffer to the additional pH values of 7.0, and 8.0. Reactions were started by addition of ATP, incubated at 37 °C for 60 min (unless otherwise indicated in figure legends) and stopped by addition of SDS-containing loading buffer. For dephosphorylation assays, after the initial

phosphorylation reaction described above, the DegS/YbdM and DegS/YabT reaction mixtures were desalted on a PD-10 column (to remove the ATP), lyophilized and resuspended in the identical reaction mixture as before, but without ATP, and incubated 2 hours at 37 °C. The proteins were separated by SDS-PAGE (for separation of PrkC and DegS we used a Tris-tricine gel). For detection of phospho-histidine and phospho-aspartate, the gels were rinsed with water and dried directly, whereas for detection of phospho-serine they were additionally treated in boiling 0.5 M HCl for 10 min. Radioactive signals were visualised using STORM phosphorImager and quantified using the ImageQuant software (GE-healthcare).

β-galactosidase assay

150 mL LB was inoculated with overnight culture to OD_{600} of 0.02 and grown with shaking at 37 °C. At time points indicated in the figures, 2 mL samples were taken, spun down (10000 g, 2 min) and cell pellets were stored at -20 °C. The pellet was resuspended in 2 mL of Z-buffer (60 mM $Na_2HPO_4{\cdot}7H_20$, 0.04 mM $NaH_2PO4{\cdot}H_20$, 10 mM KCl, 1 mM $MgSO_4$ and 50 mM β-mercaptoethanol, pH 7.0) and OD_{600} was measured. 1 mL cell suspension was treated with 0.5 mg lysozyme for 5 min at 30 °C before adding 8 µL 10 % Triton X-100 and incubating for additional 5 min. Reaction was started by addition of 100 µL of 4 mg/mL ONPG and stopped by addition of 1 mL of 0.5 M Na_2CO_3. Miller units were calculated as described previously (Miller, 1972).

Competence assays

To assess the competence of *B. subtilis* strains cells were transformed according to the two-step protocol described by Yasbin *et al*. with the modification that 60 min after dilution in GM2 medium, 0.5 mL culture were transferred to three test tubes containing 1 µg DNA and incubated an additional 90 min before plating. The plasmid pDG268-neo (Christiansen *et al*., 1997) conferring neomycin resistance was used as DNA for competence experiments. Results from three biological replicates are presented as % of competence in wild type and are average and standard deviation of three transformations. For single cell analysis the cultures were incubated for 105 min after dilution in GM2 at which time cultures were concentrated 10 fold and 5 µL were deposited on a polylysine-coated glass slide (Thermo Scientific) and examined using a Xeiss Axioplan microscope equipped with a Kappa ACC 1 condenser, a Zeiss Plan Neofluor 100x objective and a Kappa DX2 HC-FW camera. Images were acquired using Kappa Imagebase Control 2.7.2 software. In the experiments

1600 to 9800 cells per strain were examined. For wild type cells about 5.4 % were competent and values are given as % of wild type with standard deviation of two independent experiments.

Complex colony formation

Cells were grown in LB shaking at 37 °C to OD_{600} 0.5 at which point 5 µL culture was spotted on either a dried MSgg or LB plate (1.5 % agar) and incubated for 96 hours at 28 °C. Colonies were measured and photographed using a SONY Cyber-shot DSC-T20 camera with close focus enabled. For each sample, a representative image from 20 examined colonies is presented.

Swarming

Cells were grown in LB to an OD_{600} of 0.5 at which time LB plates (0.7 % agar) dried for 20 min in a fume hood (face up) were inoculated with 5 µL culture and dried an additional 10 minutes. Petri dishes were sealed with parafilm to avoid plates drying out and incubated at 37 °C. Swarm radii were measured at the times indicated in the figure. Plates were scanned about 1 hour after the swarm reached the edge of the plate using a standard HP office scanner.

Acknowledgements

We would like to thank Dan Kearns for providing the strain DS155. This work was supported by grants from the Danish National Research Council (FNU), the Lundbeckfonden and the Institut National de Recherche Agronomique (INRA) to IM and a PhD stipend from the Technical University of Denmark (DTU) to CJ.

References

Amati, G., Bisicchia, P., and Galizzi, A. (2004) DegU-P represses expression of the motility fla-che operon in Bacillus subtilis. *J Bacteriol*. 186: 6003-6014.

Amrein, K. E., Takcs, B., Stieger, M., Molnos, J., Flint, N.A., and Burn, P. (1995) Purification and characterization of recombinant human $p50^{csk}$ protein-tyrosine kinase from an *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL. *Proc Natl Acad Sci U S A* 92: 1048-1052.

Arantes, O., and Lereclus, D. (1991) Construction of cloning vectors for Bacillus thuringiensis. *Gene* 108:115-119.

Branda, S.S., González-Pastor, J.E., Ben-Yehuda, S., Losick, R., and Kolter, R. (2001) Fruiting body formation by *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 98:11621-11626.

Chang, S., and Cohen, S.N. (1979) High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. *Mol Gen Genet* 168: 111-115.

Christiansen, L.C., Schou, S., Nygaard, P., and Saxild, H.H. (1997) Xanthine Metabolism in *Bacillus subtilis*: Characterization of the *xpt-pbuX* Operon and Evidence for Purine- and Nitrogen-Controlled Expression of Genes Involved in Xanthine Salvage and Catabolism. *J Bacteriol* 179:2540–2550.

Dervyn, E., Noirot-Gros, M.-F., Mervelet, P., McGovern, S., Ehrlich, S.D., Polard, P., and Noirot, P. (2004) The bacterial condensin/cohesin-like protein complex acts in DNA repair and regulation of gene expression. *Mol Microbiol* 51:1629–1640.

Hamoen, L.W., Venema, G., and Kuipers O.P. (2003) Controlling competence in *Bacillus subtilis*: shared use of regulators. *Microbiology* 149:9-17.

Hoch, J.A. (2000) Two-component and phosphorelay signal transduction. *Curr Opin Microbiol* 3:165-170.

Laub, M.T., and Goulian, M. (2007) Specificity in two-component signal transduction pathways. *Annu Rev Genet* 41:121-145.

Jarmer, H., Berka, R., Knudsen, S., and Saxild, H.H. (2002) Transcriptome analysis documents induced competence of *Bacillus subtilis* during nitrogen limiting conditions. *FEMS Microbiol Lett* 206:197-200.

Kearns, D.B., Chu, F., Rudner, R., and Losick, R. (2004) Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. *Mol Microbiol* 52:357–369.

Kobayashi, K. (2007) Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in *Bacillus subtilis*. *Mol Microbiol* 66:395–409.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V. *et al.* (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 390:249-256.

Leonard, C.J., Aravind, L., and Koonin, E.V. (1998) Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily. *Genome Res* 8:1038-1047.

Lin, W.J., Walthers, D., Connelly, J.E., Burnside, K., Jewell, K.A., Kenney, L.J., and Rajagopal, L. (2009) Threonine phosphorylation prevents promoter DNA binding of the Group B Streptococcus response regulator CovR. *Mol Microbiol* **71**:1447-1795.

Maamar, H., and Dubnau, D. (2005) Bistability in the Bacillus subtilis K-state (competence) system requires a positive feedback loop. *Mol Microbiol* 56:615-624.

Macek, B., Mijakovic, I., Olsen, J.V., Gnad, F., Kumar, C., Jensen, P.R., and Mann, M. (2007) The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. *Mol Cell Proteomics* 6:697-707.

Madec, E., Stensballe, A., Kjellström, S., Cladière, L., Obuchowski, M., Jensen, O.N., and Séror, S.J. (2003) Mass spectrometry and site-directed mutagenesis identify several autophosphorylated residues required for the activity of PrkC, a Ser/Thr kinase from *Bacillus subtilis*. *J Mol Biol* 330:459-472.

Mäder, U., Antelmann, H., Buder, T., Dahl, M.K., Hecker, M., and Homuth, G. (2002) *Bacillus subtilis* functional genomics: genome-wide analysis of the DegS-DegU regulon by transcriptomics and proteomics. *Mol Genet Genomics* 268:455–467.

Maguin, E., Prévost, H., Ehrlich, S.D., and Gruss, A. (1996) Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. *J Bacteriol* 178:931-935.

Msadek, T., Kunst, F., Henner, D., Klier, A., Rapoport, G., and Dedonder, R. (1990) Signal transduction pathway controlling synthesis of a class of degradative enzymes in *Bacillus subtilis*: expression of the regulatory genes and analysis of mutations in degS and degU. *J Bacteriol* 172:824-834.

Mijakovic, I., Poncet, S., Boël, G., Mazé, A., Gillet, S., Jamet, E., *et al*. (2003) Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases. *EMBO J* 22:4709-4718.

Miller, J. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Mukai, K., Kawata-Mukai, M., and Tanaka, T. (1992) Stabilization of Phosphorylated *Bacillus subtilis* DegU by DegR. *J Bacteriol* 174:7954-7962.

Ogura, M., and Tanaka, T. (1997) *Bacillus subtilis* ComK negatively regulates degR gene expression. *Mol Gen Genet* 254:157-165.

Ogura, M., Yamaguchi, H., Yoshida, K.-I., Fujita, Y., and Tanaka, T. (2001) DNA microarray analysis of *Bacillus subtilis* DegU, ComA and PhoP regulons: an approach to comprehensive analysis of *B. subtilis* two-component regulatory systems. *Nucleic acids res* 29: 3804-3813.

Ogura, M., Shimane, K., Asai, K., Ogasawara, N., and Tanaka, T. (2003) Binding of response regulator DegU to the *aprE* promoter is inhibited by RapG, which is counteracted by extracellular PhrG in *Bacillus subtilis*. *Mol Microbiol* 49:1685–1697.

O'Hare, H.M., Durán, R., Cerveñansky, C., Bellinzoni, M., Wehenkel, A.M., Pritsch, O., *et al*. (2008) Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol Microbiol* 70: 1408-1423.

Petranovic, D., Grangeasse, C., Macek B., Abdillatef, M., Gueguen-Chaignon V., Nessler S., *et al*. (2009) Activation of Bacillus subtilis Ugd by the BY-Kinase PtkA Proceeds via Phosphorylation of Its Residue Tyrosine 70. *J Mol Microbiol Biotechnol* 17:83-89.

Phalip, V., Li, J.H., and Zhang, C.C. (2001) HstK, a cyanobacterial protein with both a serine/threonine kinase domain and a histidine kinase domain: implication for the mechanism of signal transduction. *Biochem J* 360:639-644.

Rasmussen, S., Nielsen, H.B., and Jarmer, H. (2009) The transcriptionally active regions in the genome of Bacillus subtilis. *Mol Microbiol* 73:1043-1057.

Shah, I.M., Laaberki, M.H., Popham, D.L., and Dworkin, J. (2008) A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. *Cell* 135:486-496.

Steil, L., Serrano, M., Henriques, A.O., and Völker, U. (2005) Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of Bacillus subtilis. *Microbiology* 151:399-420.

Ulijasz, A.T., Falk, S.P., and Weisblum, B. (2009) Phosphorylation of the RitR DNA-binding domain by a Ser-Thr phosphokinase: implications for global gene regulation in the streptococci. *Mol Microbiol* 71:382-90.

Vagner, V., Dervyn, E., and Ehrlich, S.D. (1998) A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology* 144: 3097-3104.

Verhamme, D.T., Kiley, T.B., and Stanley-Wall, N.R. (2007) DegU co-ordinates multicellular behaviour exhibited by *Bacillus subtilis*. *Mol Microbiol* 65:554–568.

Wang, S.T., Setlow, B., Conlon, E.M., Lyon, J.L., Imamura, D., Sato, T., *et al*. (2006) The forespore line of gene expression in Bacillus subtilis. *J Mol Biol* 358:16-37.

Wittekind, M., Reizer, J., Deutscher, J., Saier, M.H., and Klevit, R.E. (1989) Common structural changes accompany the functional inactivation of HPr by seryl phosphorylation or by serine to aspartate substitution. *Biochemistry* 28:9908-9912.

Yasbin, R.E., Wilson, G.A., and Young, F.E. (1975) Transformation and transfection in lysogenic strains of Bacillus subtilis: evidence for selective induction of prophage in competent cells. *J Bacteriol* 121:296-304.

Yasumura, A., Abe, S., and Tanaka, T. (2008) Involvement of Nitrogen Regulation in *Bacillus subtilis degU* Expression. *J Bacteriol* 190:5162–5171.

Østergaard, H., Henriksen, A., Hansen, F.G., and Winther, J.R. (2001) Shedding light on disulfide bond formation: engineering a redox switch in green fluorescent protein. *EMBO J* 20:5853-5862.

Chapter 4

Discussion and outlook

Discussion and outlook

Protein phosphorylation in bacteria is mediated by at least three phosphorylation systems: the two-component systems, the PTS involved in sugar uptake and finally serine/threonine and tyrosine phosphorylation. Of these systems serine/threonine and tyrosine phosphorylation are the least studied. With respect to serine/threonine phosphorylation some regulatory roles are especially well-studied exemplified by the partner switching modules in sporulation and stress response. On the other hand the Hanks type kinases in *B. subtilis*, with the exception of PrkC, remain largely uncharacterised. Tyrosine phosphorylation has been studied but research has generally been hampered by difficulties associated with identifying tyrosine phosphorylated proteins in bacteria.

In this work I took advantage of the recently published *B. subtilis* phosphoproteome that identified a number of new phosphoproteins (Macek *et al*., 2007) to study specifically tyrosine phosphorylation in relation to PtkA as well as the serine phosphorylation of DegS, the histidine kinase of the important transitional phase two component system DegS/U.

Tyrosine phosphorylation

B. subtilis encodes two proteins presently recognised as tyrosine kinases namely PtkA and the truncated BY-kinase EpsB presumed to be non-functional (Mijakovic *et al*., 2003). While the basis for concluding that EpsB is non-functional might not be convincing, the current notion is that PtkA plays a prominent role in the *B. subtilis* tyrosine phosphorylation network. For that reason we decided to test whether PtkA would phosphorylate any of the nine newly identified tyrosine phosphorylated proteins and demonstrated that it could phosphorylate all of them to various degrees *in vitro* (Chapter 2). In order to establish the proteins as true substrates of PtkA, one would need to confirm the phosphorylation *in vivo* and this is part of the current research activities in the laboratory of Mijakovic. The most important as well as most challenging task is to identify the physiological role of the phosphorylation events. The substrates were tested for PtkA-dependent phosphorylation in various assays reflecting protein function. The previously characterised substrates of PtkA were all activated by phosphorylation and in the tested conditions we identified two proteins, Asd and YorK, activated by PtkA *in vitro*. Asd is a part of the aspartate biosynthesis

pathway, but its product aspartate semialdehyde is also a precursor for diaminopimelic acid, which is a constituent of bacterial cell wall peptidoglycan linking this phosphorylation event to cell wall metabolism, a well-known role of BY-kinases. YorK is one of two RecJ homologues in *B. subtilis* (Sutera *et al*., 1999) but the protein has not been studied prior to this work where we demonstrated Mg^{2+}/Mn^{2+}-dependent exonuclease activity on single-stranded DNA as reported for RecJ in some bacteria. *Haemophilus influenzae* RecJ interacts with Ssb (Sharma and Rao, 2009), and the fact that both are substrates for PtkA in *B. subtilis* further underscores the importance of PtkA in single-stranded DNA metabolism.

Surprisingly the experiments indicated that most of the PtkA substrates were not activated by phosphorylation. We thus considered the possibility that phosphorylation might play a role in mediating protein-protein interaction and therefore the localisation in wild type and Δ*ptkA* background were examined using fluorescent protein tags. This allowed us to divide the substrates into several classes. The proteins that were activated by phosphorylation, Asd, YorK (and InfA) displayed a diffuse localisation profile. The remaining proteins showed a growth phase- and PtkA dependent localisation profile. In addition Ldh, YnfE and YvyG co-localised with PtkA while enolase and YjoA did not. In most cases the functional role of the observed localisation profiles is difficult to interpret, but in case of enolase the very strong localisation to one pole in the cell in stationary phase could indicate a role in sporulation. Enolase is a multifunctional protein that besides its role in glycolysis also is implicated in heat shock response (Miller *et al*., 1991), RNA degradation (Commichau *et al*., 2009), DNA replication (Jannière *et al*., 2007; Commichau *et al*., 2009) and is also secreted upon automodification with its substrate 2-PG (Boël *et al*., 2004). Enolase is also one of the most abundant proteins in *Bacilli* spores (DelVecchio *et al*., 2006), but whether PtkA-dependent localisation of enolase is linked to sporulation is still to be determined. The last substrate OppA is part of the Opp permease system specific for peptides of 3-5 amino acids length, where OppA is the ligand binding protein and is attached to the outside of the cell via a lipid anchor. For this protein the signal was absent in wild type cells in stationary phase while it appeared in the Δ*ptkA* background, indicating a role of PtkA in its export. As the Opp system is the sole system for uptake of tripeptides (Koide and Hoch, 1994) we tested import of peptides *in vivo*, but since it was not affected in Δ*ptkA* background, no final conclusion could be drawn in this case.

BY-kinases are organised differently in *Proteobacteria* and *Firmicutes*. In *Proteobacteria* the transmembrane modulator and the kinase domain are found on a single polypeptide, while the two

domains are separated into two polypeptides in *Firmicutes*. The solved structure of *S. aureus* BY-kinase CapB (Olivares-Illana *et al*., 2008) revealed an octameric BY-kinase ring structure anchored to the membrane *via* interaction with the transmembrane modulators. Autophosphorylation induces dissociation of the kinase monomers but the structural data suggested that kinase monomers would remain attached to the transmembrane modulators, essentially mimicking the *Proteobacteria*-like organisation. It was previously hypothesised that separation of the BY-kinase in a modulator and a cytosolic kinase might allow the *Firmicute* kinase to interact with different modulator proteins directing it to different subsets of its substrates (Mijakovic *et al*., 2005a) and this was supported by localisation data from this study. Localisation data showed that PtkA co-localised with TkmA at the membrane in exponential growth phase, while in stationary phase, it dissociated from TkmA and was found in the cytosol. We have demonstrated that PtkA ensured correct localisation of many of its substrates, of which several co-localised with PtkA, specifically in the stationary phase. It is presently unclear how localisation is ensured, whether it is phosphorylation dependent, mediated by protein-protein interaction or an indirect effect. The data do however argue for a new model in which PtkA can dissociate from its transmembrane modulator under some conditions in order to phosphorylate or otherwise interact with its substrates.

This study has significantly increased the number of known substrates of PtkA indicating roles in diverse processes such as carbon metabolism, peptide transport and motility, in addition to the previously identified roles in DNA metabolism and exopolysaccharide synthesis (figure 4.1). Further, the results indicate that BY-kinases besides phosphorylation-dependent protein activation can also act in terms of ensuring correct localisation of their substrates.

While contributing to a better understanding of tyrosine phosphorylation in *B. subtilis*, many questions remain open. In this study we have focused only on one half of the regulatory system omitting entirely the phosphatase action. The previously characterised substrates of PtkA were all targeted by PHP-type phosphatase PtpZ, and to complement this work it should be investigated whether this is also the case for the newly identified PtkA substrates. This could be extended to considering the impact of PtpZ on the localisation profile of substrates. Further, efforts should be devoted to unravelling the mechanism by which localisation is ensured. Being that PtkA is a kinase, it could seem likely that phosphorylation would be the trigger, but protein-protein interaction or some in-direct effect could also be the cause. Finally it would be of interest to analyse whether this is a conserved feature of BY-kinases or a strain-specific trait.

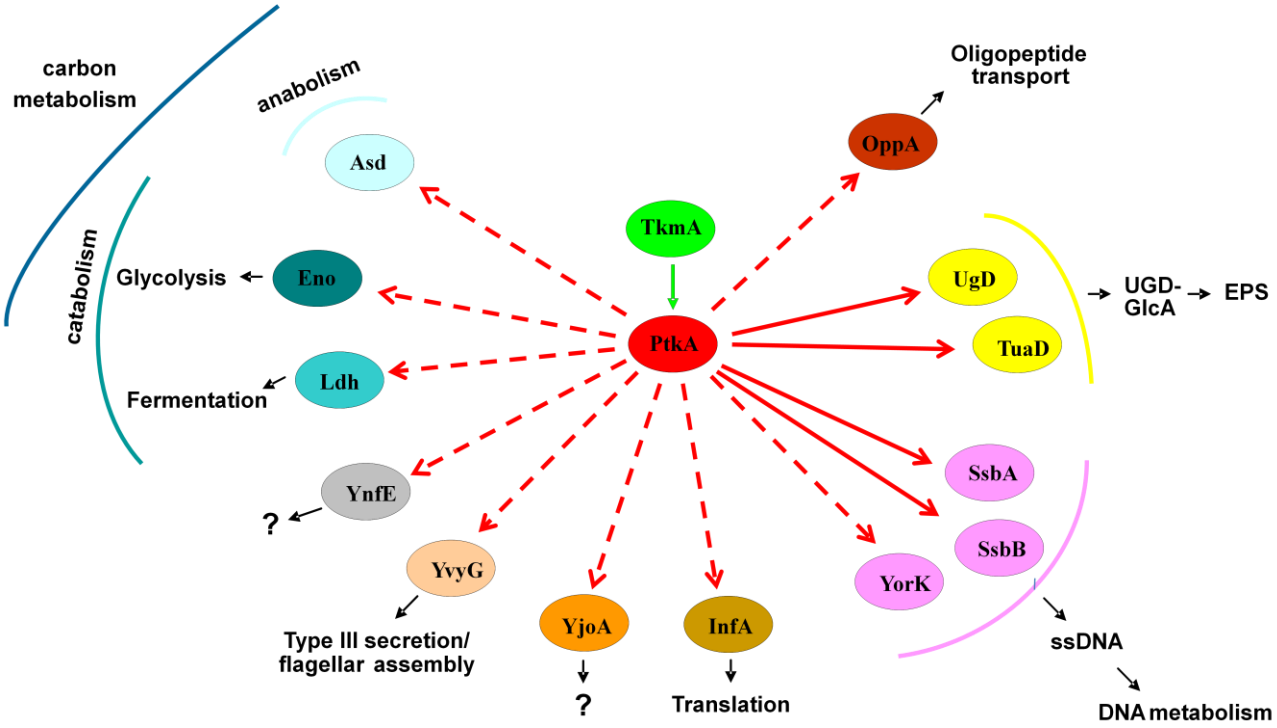


Figure 4.1: Model of the tyrosine phosphorylation network in *B. subtilis*. The kinase PtkA and modulator TkmA controls the phosphorylation and enzyme activity of previously characterised substrate classes UDP-glucose dehydrogenases and single-stranded DNA-binding proteins (bold arrows). In this work a number of new regulatory targets of PtkA were identified, most phosphorylated *in vitro* and regulated on level of enzyme activity or cellular localisation (dashed arrows).

The role of BY-kinases in cell wall metabolism and biofilm formation is well-known, and PtkA also appears to be implicated in this process via phosphorylation-dependent activation of UDP-glucose dehydrogenases. An attempt to study the role of PtkA *in vivo* in strain 168 however did not reveal a strong phenotype related to this (Petranovic *et al.*, 2007). Strain 168 is crippled in social behavioural traits such as e.g. swarming and biofilm formation, conditions under which a complex regulation could be expected to be exerted. Therefore addressing the role of phosphorylation in wild type strains such as NCIB3610 might possibly shed light on the role of PtkA in exopolysaccharide synthesis and biofilm formation and perhaps identify other regulatory roles in *B. subtilis*.

This study has identified new regulatory targets of PtkA, but it is still unclear how the kinase itself is regulated, how and when it is activated and whether, under different stimuli, it is directed against different subsets of its substrates. Interaction with the modulator protein is necessary for full activity of the kinase *in vitro* but no activating signal has so far been identified. Regulation of PtkA could also proceed via interaction with alternative modulator proteins and in a two-hybrid screen for

PtkA interactants (unpublished results) we indeed identified a potential "modulator protein" the role of which is currently being investigated.

Serine phosphorylation of two-component system kinase DegS

It is well established that DegS/U plays an important role in the transition growth phase, where it is involved in regulating phenomena such as competence, swarming, motility, biofilm formation and exoprotease production. The system has many inputs that lead to regulation on both transcriptional and post-transcriptional level. In this work the role of the recently identified phosphorylation of DegS serine-76 located in the sensing domain was studied (Chapter 3) with the working hypothesis that serine phosphorylation could be an activating signal of DegS.

DegS autophosphorylates on a histidine residue but was not subject to autophosphorylation on serine-76; instead it turned out to be phosphorylated *in vitro* by Hanks type kinases. Of the three tested kinases, YbdM and YabT were able to phosphorylate DegS but out of these two, only YbdM was specific for serine-76 and hence most likely to be the physiologically relevant kinase. With the exception of PrkC the *B. subtilis* Hanks type kinases remain largely uncharacterised. A recent study screened PrkC and YbdM for activity on eight selected substrates of which five were targeted by PrkC *in vitro* and two of these were also phosphorylated by YbdM, but to a very low degree (Pietack *et al.*, 2010). DegS, therefore, represents the first identified substrate of YbdM. The remaining characterisation was primarily done using a phospho-mimetic (S76D) and a non-phosphorylatable (S76A) version of DegS representing (theoretically) the extremes 100 and 0 % phosphorylation respectively. In line with the working hypothesis, the phospho-mimetic mutation led to an increased autophosphorylation and subsequent phosphate transfer towards DegU. DegU binds target DNA sequences in both its unphosphorylated and its phosphorylated state and at different levels of DegU phosphorylation different subregulons are activated. To confirm the regulatory potential of the phosphorylation event, we introduced genes encoding DegS S76A and S76D in place of the wild type allele and tested the effect in different physiological assays. In line with the *in vitro* data, the results pointed towards an increased pool of phosphorylated DegU in the *degS* S76D strain which exhibited a phenotype in competence and different community phenomena taking place at low and intermediate DegU~P levels. Collectively, the data point towards serine

phosphorylation being an activating signal of DegS and indicates that serine phosphorylation would correspond to an intermediary level of DegU phosphorylation *in vivo*.

The fact that mutant versions of DegS were used has the consequence that interpretation of the physiological role remains speculative. Future research should focus on confirming the kinase responsible for phosphorylation DegS *in vivo*. In this study we tried to deduce its identity by evaluating the consequence of inactivating individual kinases. Inactivating the relevant kinase would confer a DegS S76A phenotype, but our attempt failed because wild type and DegS S76A strains behaved similarly in our experimental set ups. An alternative approach could be to over-express the kinases in the hope that it would lead to a DegS S76D phenotype easily distinguishable from wild type cells. Establishing when DegS is phosphorylated and in response to what, might indicate the physiological role of the serine phosphorylation event.

This study represents the first example of a bacterial two component sensory kinase that is regulated via serine phosphorylation of its input domain by a Hanks type serine/threonine kinase. Despite the fact that DegS/U has been studied for more than 25 years, still this intriguing two-component system offers surprises.

Outlook

In this work the functional characterisation of the many newly identified phosphoproteins have been initiated providing some interesting discoveries. This is a demanding but nevertheless very important task if the improvements in phosphoproteomics are to be translated into insights in bacterial physiology. It is also of importance in relation to curating the phosphoproteomic data, as phosphopeptides, besides stemming from protein phosphorylation, can be enzyme intermediates as exemplified by e.g. phosphoglycerate mutase (Macek *et al*., 2007).

In the years to come, the breakthroughs in mass spectrometry based phosphoproteomics promise a plethora of data on protein phosphorylation. These studies will enable assignment of kinases and phosphatases to their substrates and detect differences in phosphorylation patterns in different growth conditions including in different steps in pathogenicity. This might enable the construction of bioinformatic models of global regulatory networks relying on protein phosphorylation as suggested previously (Soufi *et al*., 2008a).

This work provided an example of crosstalk between different phosphorylation systems in which a Hanks type serine/threonine kinase provided an activating signal to a two-component system sensory kinase. It is becoming clear that many of the post translational modifications thought restricted to eukaryotes are also present in bacteria as well. Recently a global study of protein lysine acetylation in *Salmonella enterica* identified 191 acetylated proteins with 90 % of the enzymes in carbon metabolism acetylated (Wang *et al*., 2010). How this meshwork of different post translational modifications work together to regulate the bacterial cell will surely keep the scientific community occupied for years to come.

References

References

Absalon, C., Obuchowski, M., Madec, E., Delattre, D., Holland, I.B., and Séror, S.J. (2009) CpgA, EF-Tu and the stressosome protein YezB are substrates of the Ser/Thr kinase/phosphatase couple, PrkC/PrpC, in *Bacillus subtilis*. *Microbiology* 155:932-943

Alper, S., Duncan, L., and Losick, R. (1994) An adenosine nucleotide switch controlling the activity of a cell type-specific transcription factor in *B. subtilis*. *Cell* 77:195-205.

Amati, G., Bisicchia, P., and Galizzi, A. (2004) DegU-P represses expression of the motility fla-che operon in *Bacillus subtilis*. *J Bacteriol*. 186: 6003-6014.

Amrein, K. E., Takcs, B., Stieger, M., Molnos, J., Flint, N.A., and Burn, P. (1995) Purification and characterization of recombinant human $p50^{csk}$ protein-tyrosine kinase from an *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL. *Proc Natl Acad Sci U S A* 92: 1048-1052.

Andersen, J.B., Roldgaard, B.B., Lindner, A.B., Christensen, B.B., and Licht, T.R. (2006) Construction of a multiple fluorescence labelling system for use in co-invasion studies of *Listeria monocytogenes*. *BMC Microbiol* 6:86.

Arantes, O., and Lereclus, D. (1991) Construction of cloning vectors for *Bacillus thuringiensis*. *Gene* 108:115-119.

Arigoni, F., Duncan, L., Alper, S., Losick, R., and Stragier, P. (1996) SpoIIE governs the phosphorylation state of a protein regulating transcription factor sigma F during sporulation in *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 93:3238-3242.

Bach, H., Wong, D., and Av-Gay, Y. (2009) *Mycobacterium tuberculosis* PtkA is a novel protein tyrosine kinase whose substrate is PtpA. *Biochem J* 420:155-160

Barford, D. (1996) Molecular mechanisms of the protein serine/threonine phosphatases. *Trends Biochem Sci* 21:407-412.

Barthe, P., Roumestand, C., Canova, M.J., Kremer, L., Hurard, C., Molle, V., and Cohen-Gonsaud, M. (2009) Dynamic and structural characterization of a bacterial FHA protein reveals a new autoinhibition mechanism. *Structure* 17:568-578.

Bendt, A.K., Burkovski, A., Schaffer, S., Bott, M., Farwick, M., and Hermann, T. (2003) Towards a phosphoproteome map of *Corynebacterium glutamicum*. *Proteomics* 3:1637-1646

Berkmen, M.B., and Grossman, A.D. (2006) Spatial and temporal organization of the *Bacillus subtilis* replication cycle. *Mol Microbiol* 62: 57-71

Boël, G., Pichereau, V., Mijakovic, I., Mazé, A., Poncet, S., Gillet, S., Giard, J.C., Hartke, A., Auffray, Y., and Deutscher, J. (2004) Is 2-phosphoglycerate-dependent automodification of bacterial enolases implicated in their export? *J Mol Biol* 337: 485-496.

Borthwick, A.C., Holms, W.H., and Nimmo, H.G. (1984) The phosphorylation of *Escherichia coli* isocitrate dehydrogenase in intact cells. *Biochem J* 222:797-804.

Branda, S.S., González-Pastor, J.E., Ben-Yehuda, S., Losick, R., and Kolter, R. (2001) Fruiting body formation by *Bacillus subtilis*. *Proc Natl Acad Sci U S A* 98:11621-11626.

Brody, M.S., Vijay, K., and Price, C.W. (2001) Catalytic function of an alpha/beta hydrolase is required for energy stress activation of the sigma(B) transcription factor in *Bacillus subtilis*. *J Bacteriol* 183:6422-6428.

Burnett, G., and Kennedy, E.P. (1954) The enzymatic phosphorylation of proteins. *J Biol Chem* 211:969-980.

Chang, S., and Cohen, S.N. (1979) High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. *Mol Gen Genet* 168: 111-115.

Chen, C.C., Lewis, R.J., Harris, R., Yudkin, M.D., and Delumeau, O. (2003) A supramolecular complex in the environmental stress signalling pathway of *Bacillus subtilis*. *Mol Microbiol* 49:1657-1669.

Chen, C.C., Yudkin, M.D., and Delumeau, O. (2004) Phosphorylation and RsbX-dependent dephosphorylation of RsbR in the RsbR-RsbS complex of *Bacillus subtilis*. *J Bacteriol* 186:6830-6836.

Christiansen, L.C., Schou, S., Nygaard, P., and Saxild, H.H. (1997) Xanthine Metabolism in *Bacillus subtilis*: Characterization of the *xpt-pbuX* Operon and Evidence for Purine- and Nitrogen-Controlled Expression of Genes Involved in Xanthine Salvage and Catabolism. *J Bacteriol* 179:2540–2550.

Collins, R.F., Beis, K., Dong, C., Botting, C.H., McDonnell, C., Ford, R.C., Clarke, B.R., Whitfield, C. and Naismith, J.H. (2007) The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *Proc Natl Acad Sci U S A*, 104:2390-2395.

Commichau, F.M., Rothe, F.M., Herzberg, C., Wagner, E., Hellwig, D., Lehnik-Habrink, M., Hammer, E., Völker, U., and Stülke, J. (2009) Novel activities of glycolytic enzymes in *Bacillus subtilis*: interactions with essential proteins involved in mRNA processing. *Mol Cell Proteomics* 8:1350-1360

Cortay, J.C., Rieul, C., Duclos, B., Cozzone, A.J. (1986) Characterization of the phosphoproteins of *Escherichia coli* cells by electrophoretic analysis. *Eur J Biochem* 159:227-237.

Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J.V. and Mann, M. (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* 4: 698-705

Cozzone, A.J., Grangeasse, C., Doublet, P., and Duclos, B. (2004) Protein phosphorylation on tyrosine in bacteria. *Arch Microbiol* 181:171-181

Dadssi, M., and Cozzone, A.J. (1990) Evidence of protein-tyrosine kinase activity in the bacterium *Acinetobacter calcoaceticus*. *J Biol Chem* 265:20996-20999.

Dufour, A., and Haldenwang, W.G. (1994) Interactions between a *Bacillus subtilis* anti-sigma factor (RsbW) and its antagonist (RsbV). *J Bacteriol* 176:1813-1820.

Dean, A.M., Lee, M.H., and Koshland, D.E. Jr. (1989) Phosphorylation inactivates *Escherichia coli* isocitrate dehydrogenase by preventing isocitrate binding. *J Biol Chem* 264:20482-20486.

Dean, A.M., and Koshland, D.E. Jr. (1990) Electrostatic and steric contributions to regulation at the active site of isocitrate dehydrogenase. *Science* 249:1044-1046.

Delvecchio, V.G., Connolly, J.P., Alefantis, T.G., Walz, A., Quan, M.A., Patra, G., Ashton, J.M., Whittington, J.T., Chafin, R.D., Liang, X., Grewal, P., Khan, A.S., and Mujer, C.V. (2006) Proteomic

profiling and identification of immunodominant spore antigens of *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*. *Appl Environ Microbiol* 72: 6355-6363

Dervyn, E., Noirot-Gros, M.-F., Mervelet, P., McGovern, S., Ehrlich, S.D., Polard, P., and Noirot, P. (2004) The bacterial condensin/cohesin-like protein complex acts in DNA repair and regulation of gene expression. *Mol Microbiol* 51:1629–1640.

Deutscher, J., Herro, R., Bourand, A., Mijakovic, I., and Poncet, S. (2005) P-Ser-HPr--a link between carbon metabolism and the virulence of some pathogenic bacteria. *Biochim Biophys Acta* 1754:118-125

Deutscher, J., Francke, C., and Postma, P.W. (2006) How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* 70:939-1031

Duncan, L., Alper, S., Arigoni, F., Losick, R., and Stragier, P. (1995) Activation of cell-specific transcription by a serine phosphatase at the site of asymmetric division. *Science* 270:641-644.

Eichenberger, P., Jensen, S.T., Conlon, E.M., van Ooij, C., Silvaggi, J., González-Pastor, J.E., Fujita, M., Ben-Yehuda, S., Stragier, P., Liu, J.S., and Losick, R. (2003) The sigmaE regulon and the identification of additional sporulation genes in *Bacillus subtilis*. *J Mol Biol* 327:945-972.

Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J., Becher, D., Gentner, S., Tam le, T., Büttner, K., Buurman, G., Scharf, C., Venz, S., Völker, U., and Hecker, M. (2004) A comprehensive proteome map of growing *Bacillus subtilis* cells. *Proteomics* 4:2849-2876.

Eymann, C., Becher, D., Bernhardt, J., Gronau, K., Klutzny, A., and Hecker, M. (2007) Dynamics of protein phosphorylation on Ser/Thr/Tyr in *Bacillus subtilis*. *Proteomics* 7:3509-3526.

Ferreira, A.S., Leitão, J.H., Sousa, S.A., Cosme, A.M., Sá-Correia, I., and Moreira, L.M. (2007) Functional analysis of *Burkholderia cepacia* genes *bceD* and *bceF*, encoding a phosphotyrosine phosphatase and a tyrosine autokinase, respectively: role in exopolysaccharide biosynthesis and biofilm formation. *Appl Environ Microbiol* 73:524-534

Fischer, E.H., and Krebs, E.G. (1955) Conversion of phosphorylase b to phosphorylase a in muscle extracts. *J Biol Chem* 216:121-132.

Fischer, C., Geourjon, C., Bourson, C., and Deutscher, J. (1996) Cloning and characterization of the *Bacillus subtilis prkA* gene encoding a novel serine protein kinase. *Gene* 168:55-60.

Fuhrmann, J., Schmidt, A., Spiess, S., Lehner, A., Turgay, K., Mechtler, K., Charpentier, E., and Clausen, T. (2009) McsB is a protein arginine kinase that phosphorylates and inhibits the heat-shock regulator CtsR. *Science* 324:1323-1327.

Fushinobu, S., Kamata, K., Iwata, S., Sakai, H., Ohta, T., and Matsuzawa, H. (1996) Allosteric activation of L-lactate dehydrogenase analyzed by hybrid enzymes with effector-sensitive and -insensitive subunits. *J Biol Chem* 271: 25611-25616

Gaidenko, T.A., Kim, T.J., and Price, C.W. (2002) The PrpC serine-threonine phosphatase and PrkC kinase have opposing physiological roles in stationary-phase *Bacillus subtilis* cells. *J Bacteriol* 184:6109-6114.

Grangeasse, C., Doublet, P., Vaganay, E., Vincent, C., Deléage, G., Duclos, B., and Cozzone, A.J. (1997) Characterization of a bacterial gene encoding an autophosphorylating protein tyrosine kinase. *Gene* 204:259-265.

Grangeasse, C., Doublet, P. and Cozzone, A.J. (2002) Tyrosine phosphorylation of protein kinase Wzc from *Escherichia coli* K12 occurs through a two-step process. *J Biol Chem* 277:7127-7135.

Grangeasse, C., Obadia, B., Mijakovic, I., Deutscher, J., Cozzone, A.J., and Doublet, P. (2003) Autophosphorylation of the *Escherichia coli* protein kinase Wzc regulates tyrosine phosphorylation of Ugd, a UDP-glucose dehydrogenase. *J Biol Chem* 278:39323-39329.

Grangeasse, C., Cozzone, A.J., Deutscher, J. and Mijakovic, I. (2007) Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology. *Trends Biochem Sci* 32:86-94.

Hahn, J., Maier, B., Haijema, B.J., Sheetz, M., and Dubnau, D. (2005) Transformation proteins and DNA uptake localize to the cell poles in *Bacillus subtilis*. *Cell* 122:59-71.

Hamoen, L.W., Venema, G., and Kuipers O.P. (2003) Controlling competence in *Bacillus subtilis*: shared use of regulators. *Microbiology* 149:9-17.

Hanks, S.K., Quinn, A.M., and Hunter, T. (1988) The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 241:42-52.

Hinc, K., Nagórska, K., Iwanicki, A., Wegrzyn, G., Séror, S.J., and Obuchowski, M. (2006) Expression of genes coding for GerA and GerK spore germination receptors is dependent on the protein phosphatase PrpE. *J Bacteriol* 188:4373-4383.

Hoch, J.A. (2000) Two-component and phosphorelay signal transduction. *Curr Opin Microbiol* 3:165-170.

Inouye, S., and Nariya, H. (2008) Dual regulation with Ser/Thr kinase cascade and a His/Asp TCS in *Myxococcus xanthus*. *Adv Exp Med Biol* 631:111-121.

Iwanicki, A., Herman-Antosiewicz, A., Pierechod, M., Séror, S.J., and Obuchowski, M. (2002) PrpE, a PPP protein phosphatase from *Bacillus subtilis* with unusual substrate specificity. *Biochem J* 366:929-936.

Iwanicki, A., Hinc, K., Seror, S., Wegrzyn, G., and Obuchowski, M. (2005) Transcription in the prpC-yloQ region in *Bacillus subtilis*. *Arch Microbiol* 183:421-430.

Jannière, L., Canceill, D., Suski, C., Kanga, S., Dalmais, B., Lestini, R., Monnier, A.F., Chapuis, J., Bolotin, A., Titok, M., Le Chatelier, E., and Ehrlich, S.D. (2007) Genetic evidence for a link between glycolysis and DNA replication. *PLoS One* 2: e447

Jarmer, H., Berka, R., Knudsen, S., and Saxild, H.H. (2002) Transcriptome analysis documents induced competence of *Bacillus subtilis* during nitrogen limiting conditions. *FEMS Microbiol Lett* 206:197-200.

Jers, C., Soufi, B., Grangeasse, C., Deutscher, J., and Mijakovic, I. (2008) Phosphoproteomics in bacteria: towards a systemic understanding of bacterial phosphorylation networks. *Expert Rev Proteomics* 5:619-627.

Kearns, D.B., Chu, F., Rudner, R., and Losick, R. (2004) Genes governing swarming in *Bacillus subtilis* and evidence for a phase variation mechanism controlling surface motility. *Mol Microbiol* 52:357–369.

Kim, T.J., Gaidenko, T.A., and Price, C.W. (2004) In vivo phosphorylation of partner switching regulators correlates with stress transmission in the environmental signaling pathway of *Bacillus subtilis*. *J Bacteriol* 186:6124-6132.

Kirstein, J., Zühlke, D., Gerth, U., Turgay, K., and Hecker, M. (2005) A tyrosine kinase and its activator control the activity of the CtsR heat shock repressor in *B. subtilis*. *EMBO J* 24:3435-3445.

Klein, G., Dartigalongue, C., and Raina, S. (2003) Phosphorylation-mediated regulation of heat shock response in *Escherichia coli*. *Mol Microbiol* 48:269-285.

Knobloch, J.K., Jäger, S., Horstkotte, M.A., Rohde, H., and Mack, D. (2004) RsbU-dependent regulation of *Staphylococcus epidermidis* biofilm formation is mediated via the alternative sigma factor sigmaB by repression of the negative regulator gene icaR. *Infect Immun* 72:3838-3848.

Kobayashi, K., Ehrlich, S.D., Albertini, A., Amati, G., Andersen, K.K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S.C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L.C., Danchin, A., Débarbouille, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S.K., Dreesen, O., Errington, J., Fillinger, S., Foster, S.J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C.R., Hecker, M., Hosoya, D., Hullo, M.F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Masson, A., Mauël, C., Meima, R., Mellado, R.P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H.M., Rapoport, G., Rawlins, J.P., Rivas, L.A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H.H., Scanlan, E., Schumann, W., Seegers, J.F., Sekiguchi, J., Sekowska, A., Séror, S.J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T., Takeuchi, M., Thomaides, H.B., Vagner, V., van Dijl, J.M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U., and Ogasawara, N. (2003) Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci U S A* 100:4678-4683.

Kobayashi, K. (2007) Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in *Bacillus subtilis*. *Mol Microbiol* 66:395–409.

Koide, A., and Hoch, J.A. (1994) Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Mol Microbiol* 13: 417-426.

Krupa, A., and Srinivasan, N. (2005) Diversity in domain architectures of Ser/Thr kinases and their homologues in prokaryotes. *BMC Genomics* 6:129.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessières, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Danchin, A. *et al.* (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 390:249-256.

Lacour, S., Bechet, E., Cozzone, A.J., Mijakovic, I., and Grangeasse, C. (2008) Tyrosine phosphorylation of the UDP-glucose dehydrogenase of *Escherichia coli* is at the crossroads of colanic acid synthesis and polymyxin resistance. *PLoS One* 3:e3053.

Lambertsen, L., Sternberg, C., and Molin, S. (2004) Mini-Tn7 transposons for site-specific tagging of bacteria with fluorescent proteins. *Environ Microbiol* 6: 726-732

LaPorte, D.C., and Chung, T. (1985) A single gene codes for the kinase and phosphatase which regulate isocitrate dehydrogenase. *J Biol Chem* 260:15291-15297.

Laub, M.T., and Goulian, M. (2007) Specificity in two-component signal transduction pathways. *Annu Rev Genet* 41:121-145.

Lee, D.C., Zheng, J., She, Y.M. and Jia, Z. (2008) Structure of *Escherichia coli* tyrosine kinase Etk reveals a novel activation mechanism. *EMBO J* 27:1758-1766.

Lee, D.C. and Jia, Z. (2009) Emerging structural insights into bacterial tyrosine kinases. *Trends Biochem Sci* 34:351-357.

Leonard, C.J., Aravind, L., and Koonin, E.V. (1998) Novel families of putative protein kinases in bacteria and archaea: evolution of the "eukaryotic" protein kinase superfamily. *Genome Res* 8:1038-1047.

Lévine, A., Vannier, F., Absalon, C., Kuhn, L., Jackson, P., Scrivener, E., Labas, V., Vinh, J., Courtney, P., Garin, J., and Séror, S.J. (2006) Analysis of the dynamic *Bacillus subtilis* Ser/Thr/Tyr phosphoproteome implicated in a wide variety of cellular processes. *Proteomics* 6:2157-2173

Lin, M.H., Hsu, T.L., Lin, S.Y., Pan, Y.J., Jan, J.T., Wang, J.T., Khoo, K.H., and Wu, S.H. (2009a) Phosphoproteomics of *Klebsiella pneumoniae* NTUH-K2044 reveals a tight link between tyrosine phosphorylation and virulence. *Mol Cell Proteomics* 8:2613-2623

Lin, W.J., Walthers, D., Connelly, J.E., Burnside, K., Jewell, K.A., Kenney, L.J., Rajagopal, L. (2009b) Threonine phosphorylation prevents promoter DNA binding of the Group B *Streptococcus* response regulator CovR. *Mol Microbiol* 71:1477-1495.

Lindner, C., Nijland, R., van Hartskamp, M., Bron, S., Hamoen, L.W., and Kuipers, O.P. (2004) Differential expression of two paralogous genes of *Bacillus subtilis* encoding single-stranded DNA binding protein. *J Bacteriol* 186:1097-1105.

Lu, T., Tan, H., Lee, D., Chen, G. and Jia, Z. (2009) New insights into the activation of *Escherichia coli* tyrosine kinase revealed by molecular dynamics simulation and biochemical analysis. *Biochemistry* 48:7986-7995.

Lukong, K.E., Larocque, D., Tyner, A.L., and Richard, S. (2005) Tyrosine phosphorylation of sam68 by breast tumor kinase regulates intranuclear localization and cell cycle progression. *J Biol Chem* 280: 38639-38647

Maamar, H., and Dubnau, D. (2005) Bistability in the *Bacillus subtilis* K-state (competence) system requires a positive feedback loop. *Mol Microbiol* 56:615-624.

Macek, B., Mijakovic, I., Olsen, J.V., Gnad, F., Kumar, C., Jensen, P.R., and Mann, M. (2007) The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. *Mol Cell Proteomics* 6:697-707.

Macek, B., Gnad, F., Soufi, B., Kumar, C., Olsen, J.V., Mijakovic, I., and Mann, M. (2008) Phosphoproteome analysis of *E. coli* reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Mol Cell Proteomics* 7:299-307.

Madec, E., Laszkiewicz, A., Iwanicki, A., Obuchowski, M., and Séror, S. (2002) Characterization of a membrane-linked Ser/Thr protein kinase in *Bacillus subtilis*, implicated in developmental processes. *Mol Microbiol* 46:571-586.

Madec, E., Stensballe, A., Kjellström, S., Cladière, L., Obuchowski, M., Jensen, O.N., and Séror, S.J. (2003) Mass spectrometry and site-directed mutagenesis identify several autophosphorylated residues required for the activity of PrkC, a Ser/Thr kinase from *Bacillus subtilis*. *J Mol Biol* 330:459-472.

Madeo, F., Schlauer, J., Zischka, H., Mecke, D., and Fröhlich, K.U. (1998) Tyrosine phosphorylation regulates cell cycle-dependent nuclear localization of Cdc48p. *Mol Biol Cell* 9: 131-141

Mäder, U., Antelmann, H., Buder, T., Dahl, M.K., Hecker, M., and Homuth, G. (2002) *Bacillus subtilis* functional genomics: genome-wide analysis of the DegS-DegU regulon by transcriptomics and proteomics. *Mol Genet Genomics* 268:455–467.

Maguin, E., Prévost, H., Ehrlich, S.D., and Gruss, A. (1996) Efficient insertional mutagenesis in lactococci and other gram-positive bacteria. *J Bacteriol* 178:931-935.

Marchler-Bauer, A., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., He, S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A., Liu, C., Lu, F., Lu, S., Marchler, G.H., Mullokandov, M., Song, J.S., Tasneem, A., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., and Bryant, S.H. (2009) CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* 37(Database issue):D205-210.

Martin-Verstraete, I., Débarbouillé, M., Klier, A., and Rapoport, G. (1990) Levanase Operon of *Bacillus subtilis* Includes a Fructose-specific Phosphotransferase System Regulating the expression of the Operon. *J. mol. biol* 214: 657-671

Martin-Verstraete, I., Deutscher, J., and Galinier, A. (1999) Phosphorylation of HPr and Crh by HprK, early steps in the catabolite repression signalling pathway for the *Bacillus subtilis* levanase operon. *J Bacteriol* 181:2966-2969.

Mascher, T., Helmann, J.D., and Unden, G. (2006) Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol Mol Biol Rev* 70:910-938

Masuda, S., Murakami, K.S., Wang, S., Anders Olson, C., Donigian, J., Leon, F., Darst, S.A., and Campbell, E.A. (2004) Crystal structures of the ADP and ATP bound forms of the *Bacillus* anti-sigma factor SpoIIAB in complex with the anti-anti-sigma SpoIIAA. *J Mol Biol* 340:941-956.

Mattoo, S., Yuk, M.H., Huang, L.L., and Miller, J.F. (2004) Regulation of type III secretion in *Bordetella*. *Mol Microbiol* 52:1201-1214.

Meile, J.C., Wu, L.J., Ehrlich, S.D., Errington, J., and Noirot, P. (2006) Systematic localisation of proteins fused to the green fluorescent protein in *Bacillus subtilis*: identification of new proteins at the DNA replication factory. *Proteomics* 6: 2135-2146

Mijakovic, I., Poncet, S., Galinier, A., Monedero, V., Fieulaine, S., Janin, J., Nessler, S., Marquez, J.A., Scheffzek, K., Hasenbein, S., Hengstenberg, W., and Deutscher, J. (2002) Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: a relic of early life? *Proc Natl Acad Sci U S A* 99:13442-13447.

Mijakovic, I., Poncet, S., Boël, G., Mazé, A., Gillet, S., Jamet, E., Decottignies, P., Grangeasse, C., Doublet, P., Le Maréchal, P., and Deutscher, J. (2003) Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases. *EMBO J* 22:4709-4718.

Mijakovic, I., Petranovic, D., and Deutscher, J. (2004) How tyrosine phosphorylation affects the UDP-glucose dehydrogenase activity of *Bacillus subtilis* YwqF. *J Mol Microbiol Biotechnol* 8: 19-25

Mijakovic, I., Petranovic, D., Bottini, N., Deutscher, J., and Jensen, P.R. (2005a) Protein-tyrosine phosphorylation in *Bacillus subtilis*. *J Mol Microbiol Biotechnol* 9:189-197.

Mijakovic, I., Musumeci, L., Tautz, L., Petranovic, D., Edwards, R.A., Jensen, P.R., Mustelin, T., Deutscher, J., and Bottini, N. (2005b) In vitro characterization of the *Bacillus subtilis* protein tyrosine phosphatase YwqE. *J Bacteriol* 187:3384-3390.

Mijakovic, I., Petranovic, D., Macek, B., Cepo, T., Mann, M., Davies, J., Jensen, P.R., and Vujaklija, D. (2006) Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res* 34:1588-1596.

Miller, B.S., Kennedy, T.E., and Streips, U.N. (1991) Molecular characterization of specific heat shock proteins in *Bacillus subtilis*. *Curr Microbiol* 22: 231-236

Miller, J. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Miller, M.L., Soufi, B., Jers, C., Blom, N., Macek, B., and Mijakovic, I. (2009) NetPhosBac - a predictor for Ser/Thr phosphorylation sites in bacterial proteins. *Proteomics* 9:116-125.

Min, K.T., Hilditch, C.M., Diederich, B., Errington, J., and Yudkin, M.D. (1993) Sigma F, the first compartment-specific transcription factor of *B. subtilis*, is regulated by an anti-sigma factor that is also a protein kinase. *Cell* 74:735-742.

Minic, Z., Marie, C., Delorme, C., Faurie, J.M., Mercier, G., Ehrlich, D., and Renault, P. (2007) Control of EpsE, the phosphoglycosyltransferase initiating exopolysaccharide synthesis in *Streptococcus thermophilus*, by EpsD tyrosine kinase. *J Bacteriol* 189:1351-1357

Morona, J.K., Morona, R., Miller, D.C., and Paton, J.C. (2002) *Streptococcus pneumoniae* capsule biosynthesis protein CpsB is a novel manganese-dependent phosphotyrosine-protein phosphatase. *J Bacteriol* 184:577-583

Morona, J.K., Morona, R., Miller, D.C., and Paton, J.C. (2003) Mutational analysis of the carboxy-terminal (YGX)4 repeat domain of CpsD, an autophosphorylating tyrosine kinase required for capsule biosynthesis in *Streptococcus pneumoniae*. *J Bacteriol* 185:3009-3019

Morona, J.K., Miller, D.C., Morona, R., and Paton, J.C. (2004) The effect that mutations in the conserved capsular polysaccharide biosynthesis genes cpsA, cpsB, and cpsD have on virulence of *Streptococcus pneumoniae*. *J Infect Dis* 189):1905-1913

Mougous, J.D., Gifford, C.A., Ramsdell, T.L., and Mekalanos, J.J. (2007) Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aeruginosa*. *Nat Cell Biol* 9:797-803.

Msadek, T., Kunst, F., Henner, D., Klier, A., Rapoport, G., and Dedonder, R. (1990) Signal transduction pathway controlling synthesis of a class of degradative enzymes in *Bacillus subtilis*: expression of the regulatory genes and analysis of mutations in degS and degU. *J Bacteriol* 172:824-834.

Mukai, K., Kawata-Mukai, M., and Tanaka, T. (1992) Stabilization of Phosphorylated *Bacillus subtilis* DegU by DegR. *J Bacteriol* 174:7954-7962.

Muñoz-Dorado, J., Inouye, S., and Inouye, M. (1991) A gene encoding a protein serine/threonine kinase is required for normal development of *M. xanthus*, a gram-negative bacterium. *Cell* 67:995-1006.

Murray, E.J., Kiley, T.B., and Stanley-Wall, N.R. (2009) A pivotal role for the response regulator DegU in controlling multicellular behaviour. *Microbiology* 155:1-8.

Musumeci, L., Bongiorni, C., Tautz, L., Edwards, R.A., Osterman, A., Perego, M., Mustelin, T., and Bottini, N. (2005) Low-molecular-weight protein tyrosine phosphatases of *Bacillus subtilis*. *J Bacteriol* 187:4945-4956.

Nariya, H., and Inouye, S. (2005) Identification of a protein Ser/Thr kinase cascade that regulates essential transcriptional activators in *Myxococcus xanthus* development. *Mol Microbiol* 58:367-379.

Obadia, B., Lacour, S., Doublet, P., Baubichon-Cortay, H., Cozzone, A.J., and Grangeasse, C. (2007) Influence of tyrosine-kinase Wzc activity on colanic acid production in *Escherichia coli* K12 cells. *J Mol Biol* 367:42-53.

Ogura, M., and Tanaka, T. (1997) *Bacillus subtilis* ComK negatively regulates degR gene expression. *Mol Gen Genet* 254:157-165.

Ogura, M., Yamaguchi, H., Yoshida, K.-I., Fujita, Y., and Tanaka, T. (2001) DNA microarray analysis of *Bacillus subtilis* DegU, ComA and PhoP regulons: an approach to comprehensive analysis of *B. subtilis* two-component regulatory systems. *Nucleic acids res* 29: 3804-3813.

Ogura, M., Shimane, K., Asai, K., Ogasawara, N., and Tanaka, T. (2003) Binding of response regulator DegU to the *aprE* promoter is inhibited by RapG, which is counteracted by extracellular PhrG in *Bacillus subtilis*. *Mol Microbiol* 49:1685–1697.

O'Hare, H.M., Durán, R., Cerveñansky, C., Bellinzoni, M., Wehenkel, A.M., Pritsch, O., Obal, G., Baumgartner, J., Vialaret, J., Johnsson, K., and Alzari, P.M. (2008) Regulation of glutamate metabolism by protein kinases in mycobacteria. *Mol Microbiol* 70: 1408-1423.

Olivares-Illana, V., Meyer, P., Bechet, E., Gueguen-Chaignon, V., Soulat, D., Lazereg-Riquier, S., Mijakovic, I., Deutscher, J., Cozzone, A.J., Laprévote, O., Morera, S., Grangeasse, C. and Nessler, S. (2008) Structural basis for the regulation mechanism of the tyrosine kinase CapB from *Staphylococcus aureus*. *PLoS Biol* 6:e143.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1:376-386.

Pagni, M., Lazarevic, V., Soldo, B., and Karamata, D. (1999) Assay for UDPglucose 6-dehydrogenase in phosphate-starved cells: gene tuaD of *Bacillus subtilis* 168 encodes the UDPglucose 6-dehydrogenase involved in teichuronic acid synthesis. *Microbiology* 145:1049-1053.

Paiment, A., Hocking, J., and Whitfield, C. (2002) Impact of phosphorylation of specific residues in the tyrosine autokinase, Wzc, on its activity in assembly of group 1 capsules in *Escherichia coli*. *J Bacteriol* 184:6437-6447.

Pallen, M., Chaudhuri, R., and Khan, A. (2002) Bacterial FHA domains: neglected players in the phospho-threonine signalling game? *Trends Microbiol* 10:556-563.

Pallen, M.J., Penn, C.W., and Chaudhuri, R.R. (2005) Bacterial flagellar diversity in the post-genomic era. *Trends Microbiol* 13:143-149.

Peleg, A., Shifrin, Y., Ilan, O., Nadler-Yona, C., Nov, S., Koby, S., Baruch, K., Altuvia, S., Elgrably-Weiss, M., Abe, C.M., Knutton, S., Saper, M.A., and Rosenshine, I. (2005) Identification of an *Escherichia coli* operon required for formation of the O-antigen capsule. *J Bacteriol* 187:5259-5266.

Petranovic, D., Michelsen, O., Zahradka, K., Silva, C., Petranovic, M., Jensen, P.R., and Mijakovic, I. (2007) *Bacillus subtilis* strain deficient for the protein-tyrosine kinase PtkA exhibits impaired DNA replication. *Mol Microbiol* 63:1797-1805.

Petranovic, D., Grangeasse, C., Macek, B., Abdillatef, M., Gueguen-Chaignon, V., Nessler, S., Deutscher, J., and Mijakovic, I. (2009) Activation of *Bacillus subtilis* Ugd by the BY-kinase PtkA proceeds via phosphorylation of its residue tyrosine 70. *J Mol Microbiol Biotechnol* 17:83-89

Phalip, V., Li, J.H., and Zhang, C.C. (2001) HstK, a cyanobacterial protein with both a serine/threonine kinase domain and a histidine kinase domain: implication for the mechanism of signal transduction. *Biochem J* 360:639-644.

Pietack, N., Becher, D., Schmidl, S.R., Saier, M.H., Hecker, M., Commichau, F.M., and Stülke, J. (2010) In vitro Phosphorylation of Key Metabolic Enzymes from *Bacillus subtilis*: PrkC Phosphorylates Enzymes from Different Branches of Basic Metabolism. *J Mol Microbiol Biotechnol* 18:129-140.

Ptacek, J., Devgan, G., Michaud, G., Zhu, H., Zhu, X., Fasolo, J., Guo, H., Jona, G., Breitkreutz, A., Sopko, R., McCartney, R.R., Schmidt, M.C., Rachidi, N., Lee, S.J., Mah, A.S., Meng, L., Stark, M.J., Stern, D.F., De Virgilio, C., Tyers, M., Andrews, B., Gerstein, M., Schweitzer, B., Predki, P.F., and Snyder, M. (2005) Global analysis of protein phosphorylation in yeast. *Nature* 438: 679-684

Rappsilber, J., Mann, M. and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2: 1896-1906

Rasmussen, S., Nielsen, H.B., and Jarmer, H. (2009) The transcriptionally active regions in the genome of *Bacillus subtilis*. *Mol Microbiol* 73:1043-1057.

Ravichandran, A., Sugiyama, N., Tomita, M., Swarup, S., and Ishihama, Y. (2009) Ser/Thr/Tyr phosphoproteome analysis of pathogenic and non-pathogenic *Pseudomonas* species. *Proteomics* 9:2764-2775.

Romero, S., Merino, E., Bolívar, F., Gosset, G., and Martinez, A. (2007) Metabolic engineering of *Bacillus subtilis* for ethanol production: lactate dehydrogenase plays a key role in fermentative metabolism. *Appl Environ Microbiol* 73: 5190-5198

Rubin, C.S., and Rosen, O.M. (1975) Protein phosphorylation. *Annu Rev Biochem* 44:831-887

Sandler, N., and Keynan, A. (2007) Cell cycle dependent phosphatase activity in *Bacillus subtilis*. *J Basic Microbiol* 47:222-229.

Schmidl, S.R., Gronau, K., Pietack, N., Hecker, M., Becher, D., and Stülke, J. (2010) The phosphoproteome of the minimal bacterium *Mycoplasma pneumoniae*: Analysis of the complete known Ser/ Thr kinome suggests the existence of novel kinases. *Mol Cell Proteomics* 9:1228-1242.

Shah, I.M., Laaberki, M.H., Popham, D.L., and Dworkin, J. (2008) A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. *Cell* 135:486-496.

Shah, I.M., and Dworkin, J. (2010) Induction and regulation of a secreted peptidoglycan hydrolase by a membrane Ser/Thr kinase that detects muropeptides. *Mol Microbiol* 75:1232-1243.

Sharma, R., and Rao, D.N. (2009) Orchestration of *Haemophilus influenzae* RecJ exonuclease by interaction with single-stranded DNA-binding protein. *J Mol Biol* 385: 1375-1396

Shi, L., Kobir, A., Jers, C. and Mijakovic, I. (2010) Bacterial protein-tyrosine kinases. *Curr proteomics* (in press)

Singh, K.D., Halbedel, S., Görke, B., and Stülke, J. (2007) Control of the phosphorylation state of the HPr protein of the phosphotransferase system in *Bacillus subtilis*: implication of the protein phosphatase PrpC. *J Mol Microbiol Biotechnol* 13:165-171.

Soufi, B., Jers, C., Hansen, M.E., Petranovic, D., and Mijakovic, I. (2008a) Insights from site-specific phosphoproteomics in bacteria. *Biochim Biophys Acta* 1784:186-192.

Soufi, B., Gnad, F., Jensen, P.R., Petranovic, D., Mann, M., Mijakovic, I., and Macek, B. (2008b) The Ser/Thr/Tyr phosphoproteome of *Lactococcus lactis* IL1403 reveals multiply phosphorylated proteins. *Proteomics* 8:3486-3493.

Soulat, D., Jault, J.M., Duclos, B., Geourjon, C., Cozzone, A.J., and Grangeasse, C. (2006) *Staphylococcus aureus* operates protein-tyrosine phosphorylation through a specific mechanism. *J Biol Chem* 281:14048-14056

Soulat, D., Grangeasse, C., Vaganay, E., Cozzone, A.J., and Duclos, B. (2007) UDP-acetyl-mannosamine dehydrogenase is an endogenous protein substrate of *Staphylococcus aureus* protein-tyrosine kinase activity. *J Mol Microbiol Biotechnol* 13:45-54

Soung, G.Y., Miller, J.L., Koc, H., and Koc, E.C. (2009) Comprehensive analysis of phosphorylated proteins of *Escherichia coli* ribosomes. *J Proteome Res* 8:3390-3402.

Steil, L., Serrano, M., Henriques, A.O., and Völker, U. (2005) Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of *Bacillus subtilis*. *Microbiology* 151:399-420.

Stock, A.M., Robinson, V.L., and Goudreau, P.N. (2000) Two-component signal transduction. *Annu Rev Biochem* 69:183-215.

Sutera, V.A. Jr, Han, E.S., Rajman, L.A., and Lovett, S.T. (1999) Mutational analysis of the RecJ exonuclease of *Escherichia coli*: identification of phosphoesterase motifs. *J Bacteriol* 181: 6098-6102

Thomasson, B., Link, J., Stassinopoulos, A.G., Burke, N., Plamann, L., and Hartzell, P.L. (2002) MglA, a small GTPase, interacts with a tyrosine kinase to control type IV pili-mediated motility and development of *Myxococcus xanthus*. *Mol Microbiol* 46:1399-1413.

Ulijasz, A.T., Falk, S.P., and Weisblum, B. (2009) Phosphorylation of the RitR DNA-binding domain by a Ser-Thr phosphokinase: implications for global gene regulation in the streptococci. *Mol Microbiol* 71:382-390.

Vagner, V., Dervyn, E., and Ehrlich, S.D. (1998) A vector for systematic gene inactivation in *Bacillus subtilis*. *Microbiology*. 144: 3097-3104.

Veening, J.W., Smits, W.K., Hamoen, L.W., Jongbloed, J.D., and Kuipers, O.P. (2004) Visualization of differential gene expression by improved cyan fluorescent protein and yellow fluorescent protein production in *Bacillus subtilis*. *Appl Environ Microbiol* 70: 6809-6815

Verhamme, D.T., Kiley, T.B., and Stanley-Wall, N.R. (2007) DegU co-ordinates multicellular behaviour exhibited by *Bacillus subtilis*. *Mol Microbiol* 65:554–568.

Vijay, K., Brody, M.S., Fredlund, E., and Price, C.W. (2000) A PP2C phosphatase containing a PAS domain is required to convey signals of energy stress to the sigmaB transcription factor of *Bacillus subtilis*. *Mol Microbiol* 35:180-188.

Vincent, C., Doublet, P., Grangeasse, C., Vaganay, E., Cozzone, A.J., and Duclos, B. (1999) Cells of *Escherichia coli* contain a protein-tyrosine kinase, Wzc, and a phosphotyrosine-protein phosphatase, Wzb. *J Bacteriol* 181: 3472-3477

Vincent, C., Duclos, B., Grangeasse, C., Vaganay, E., Riberty, M., Cozzone, A.J., and Doublet, P. (2000) Relationship between exopolysaccharide production and protein-tyrosine phosphorylation in gram-negative bacteria. *J Mol Biol* 304:311-321.

Voigt, B., Antelmann, H., Albrecht, D., Ehrenreich, A., Maurer, K.H., Evers, S., Gottschalk, G., van Dijl, J.M., Schweder, T., and Hecker, M. (2009) Cell physiology and protein secretion of Bacillus licheniformis compared to *Bacillus subtilis*. *J Mol Microbiol Biotechnol* 16:53-68.

Voisin, S., Watson, D.C., Tessier, L., Ding, W., Foote, S., Bhatia, S., Kelly, J.F., and Young, N.M. (2007) The cytoplasmic phosphoproteome of the Gram-negative bacterium *Campylobacter jejuni*: evidence for modification by unidentified protein kinases. *Proteomics* 7:4338-4348

Wang, J.Y., and Koshland, D.E. Jr. (1978) Evidence for protein kinase activities in the prokaryote *Salmonella typhimurium*. *J Biol Chem* 253:7605-7608.

Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y., Ning, Z.B., Zeng, R., Xiong, Y., Guan, K.L., Zhao, S., Zhao, G.P. (2010) Acetylation of metabolic enzymes coordinates carbon source utilization and metabolic flux. *Science* 327:1004-1007.

Wang, S.T., Setlow, B., Conlon, E.M., Lyon, J.L., Imamura, D., Sato, T., Setlow, P., Losick, R., and Eichenberger, P. (2006) The forespore line of gene expression in *Bacillus subtilis*. *J Mol Biol* 358:16-37.

Whitfield, C. (2006) Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu Rev Biochem* 75: 39-68

Wittekind, M., Reizer, J., Deutscher, J., Saier, M.H., and Klevit, R.E. (1989) Common structural changes accompany the functional inactivation of HPr by seryl phosphorylation or by serine to aspartate substitution. *Biochemistry* 28:9908-9912.

Wolff, S., Otto, A., Albrecht, D., Zeng, J.S., Büttner, K., Glückmann, M., Hecker, M., and Becher, D. (2006) Gel-free and gel-based proteomics in *Bacillus subtilis*: a comparative study. *Mol Cell Proteomics* 5:1183-1192.

Wu, J., Ohta, N., Zhao, J.L., and Newton, A. (1999) A novel bacterial tyrosine kinase essential for cell division and differentiation. *Proc Natl Acad Sci U S A* 96:13068-13073.

Yang, X., Kang, C.M., Brody, M.S., and Price C.W. (1996) Opposing pairs of serine protein kinases and phosphatases transmit signals of environmental stress to activate a bacterial transcription factor. *Genes Dev* 10:2265-2275.

Yasbin, R.E., Wilson, G.A., and Young, F.E. (1975) Transformation and transfection in lysogenic strains of *Bacillus subtilis*: evidence for selective induction of prophage in competent cells. *J Bacteriol* 121:296-304.

Yasumura, A., Abe, S., and Tanaka, T. (2008) Involvement of Nitrogen Regulation in *Bacillus subtilis degU* Expression. *J Bacteriol* 190:5162–5171.

Zhao, X., and Lam, J.S. (2002) WaaP of *Pseudomonas aeruginosa* is a novel eukaryotic type protein-tyrosine kinase as well as a sugar kinase essential for the biosynthesis of core lipopolysaccharide. *J Biol Chem* 277:4722-4730.

Østergaard, H., Henriksen, A., Hansen, F.G., and Winther, J.R. (2001) Shedding light on disulfide bond formation: engineering a redox switch in green fluorescent protein. *EMBO J* 20:5853-5862

.

List of appendices

I. Soufi, B., Jers, C., Hansen, M.E., Petranovic, D. and Mijakovic, I. (2008) Insights gained from bacterial phosphoproteomics. *Biochim Biophys Acta* 1784:186-192

II. Jers, C., Soufi, B., Grangeasse, C., Deutscher, J., and Mijakovic, I. (2008) Phosphoproteomics in bacteria: towards a systemic understanding of bacterial phosphorylation networks. *Expert Rev Proteomics* 5:619-627

III. Miller, M.L., Soufi, B., Jers, C., Blom, N., Macek, B., and Mijakovic, I. (2009) NetPhosBac - a predictor for Ser/Thr phosphorylation sites in bacterial proteins. *Proteomics* 9:116-125

IV. Shi, L., Kobir, A., Jers, C. and Mijakovic, I. (2010) Bacterial protein-tyrosine kinases. *Current proteomics* (In press)

V. Co-author statements

Available online at www.sciencedirect.com
ScienceDirect
Biochimica et Biophysica Acta 1784 (2008) 186–192
Biochimica et Biophysica Acta
BBA
www.elsevier.com/locate/bbapap

Review

Insights from site-specific phosphoproteomics in bacteria

Boumediene Soufi, Carsten Jers, Mette Erichsen Hansen, Dina Petranovic, Ivan Mijakovic *

Center for Microbial Biotechnology, BioCentrum, Technical University of Denmark, Building 301, DK-2800 Lyngby, Denmark

Received 19 June 2007; received in revised form 26 July 2007; accepted 30 July 2007
Available online 15 August 2007

Abstract

Recent advances in mass spectrometry allowed the charting of bacterial serine/threonine/tyrosine phosphoproteomes with unprecedented accuracy, including the acquisition of a large number of phosphorylation sites. Phosphorylated bacterial proteins are involved in some key housekeeping processes, and their phosphorylation is expected to play an important regulatory role. When coupled to stable isotope labeling by amino acids in cell culture (SILAC), high-resolution mass spectrometry allows the detection of changes in the occupancy of phosphorylation sites in response to various stimuli. This and similar approaches promise to lead bacterial phosphoproteomics into the era of systems biology, where the entire phosphorylation-based regulatory networks will be charted, modelled, and ultimately engineered to obtain desired properties.

Keywords: Bacterial phosphoproteome; Phosphoproteomics; Protein phosphorylation; Kinases; Phosphatases; Phosphopredictor

1. Historical overview of phosphoproteomics in bacteria

Protein phosphorylation is a post-translational modification that is widely recognised as an important means of signal transduction and regulation in the living cell. Protein phosphorylation on serine/threonine/tyrosine residues is considered to be predominant in eukaryal cells [1], whereas phosphorylation on histidine/aspartate residues is usually associated with bacterial signalling [2]. In fact, it was not until the end of the 1970s that isocitrate dehydrogenase (IDH), a tricarboxylic acid cycle enzyme from *Escherichia coli*, became the first bacterial protein known to be phosphorylated on a serine residue [3]. This enzyme was found to be phosphorylated on a serine residue by a kinase which was later also shown to dephosphorylate phospho-IDH [4]. Surprisingly, no sequence similarity was found when comparing IDH kinase/phosphatase with eukaryotic protein kinases or phosphatases. Soon thereafter, a second bacterial protein was found to be phosphorylated on serine: the phosphocarrier protein HPr, from the phosphoenolpyruvate-dependant carbohydrate phosphotransferase system of *Streptococcus pyogenes* [5,6]. Interestingly, the kinase phosphorylating HPr also turned out to be bifunctional [7]; it was named HPr kinase/phosphorylase because the dephosphorylation event is a phosphorolysis reaction instead of hydrolysis [8]. In 1986, Cortay and coworkers showed that phosphorylation of bacterial proteins also could take place on a threonine or a tyrosine residue [9] and in 1996, the very first bacterial protein-tyrosine kinase called Ptk was purified from *Acinetobactor johnsonii* [10]. Ptk autophosphorylated at the expense of ATP on multiple tyrosine residues. However, like IDH kinase/phosphatase and HPr kinase/phosphorylase, Ptk did not structurally resemble any of the known eukaryal protein kinases. Indeed, most bacterial tyrosine kinases known today are classified in a separate functional group of BY-kinases [11]. In addition to autophosphorylation, BY-kinases have been found to phosphorylate, and thereby regulate the activity of many endogenous proteins [12–16], affecting cellular processes such as exoploysaccharide synthesis, antibiotic resistance, DNA metabolism and stress response. Bacteria also possess eukaryal-like serine/threonine kinases, many of which have been characterized and shown to participate in specific regulatory mechanisms [17–19]. Bacterial pathogens also have been found to use protein phosphorylation quite extensively in

Abbreviations: BY-kinase, Bacterial tyrosine kinase; GPS, group-based phosphorylation scoring; HPr, histidine phosphocarrier protein; MS, mass spectrometry; PTS, phosphoenolpyruvate-dependant carbohydrate phosphotransferase system; SILAC, stable isotope labeling by amino acids in cell culture

* Corresponding author. Tel.: +45 45 25 24 95; fax: +45 45 93 28 09.
E-mail address: im@biocentrum.dtu.dk (I. Mijakovic).

1570-9639/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2007.07.018

biochemical warfare, aimed at impairing the host defences through interference with its phosphorylation-dependent signalling pathways [20].

Shortly after the first reports on protein phosphorylation on serine/threonine/tyrosine in bacteria, evidence for this type of protein phosphorylation emerged in archaea. Serine and threonine phosphorylated proteins were first found in *Halobacterium halobium* (now *H. salinarium*) [21] and *Sulfolobus acidocaldarius* [22], whereas tyrosine phosphorylation was first detected in *Thermococcus kodakaraensis* [23]. However, compared to bacteria, only a small number of archaeal phosphoproteins has been identified, and the information on the functional aspects of phosphorylation is still quite scarce [24,25]. In 2004, a very useful database was set up, listing all bacterial and archaeal phosphoproteins characterized to date [26]. This regularly updated phosphorylation site database provides access to information on over one hundred serine/threonine/tyrosine phosphorylated proteins, making all relevant data from primary publications easily accessible. Since the site-specific data were only available from individual biochemical studies of bacterial phosphoproteins, the database is mainly built up on such studies.

Phosphoproteomics, focusing on the global studies of protein phosphorylation, commenced in bacteria with the development of 2D-gel electrophoresis. In *E. coli*, early phosphoproteomic studies with ^{32}P-radiolabelling indicated the presence of over 100 proteins phosphorylated at serine/threonine/tyrosine residues in this bacterial cell [9]. Initially, identification of phosphoproteins was accomplished by Edman degradation, and it was subsequently facilitated with the use of mass spectrometry. The resolution of protein on 2D-gels made it possible to identify these phospho-proteins [27], and even study the dynamics of bacterial protein phosphorylation in different growth conditions [28]. One phosphoproteomic study suggested that polyphosphorylated proteins accumulate in the bacterial cell under stress conditions, and thereby tag the affected proteins for degradation [29]. These 2D gel-based approaches provided many useful insights, but fell short of determining the exact phosphorylation sites in bacterial phosphoproteomes. In 2007, this finally became possible due to the advent of high resolution/high accuracy mass spectrometry systems capable of resolving mixtures as complex as entire cellular proteomes [30], which resulted in the first site-specific study of a bacterial phosphoproteome [31]. This study identified over 100 phosphorylated peptides, and pin-pointed 78 phosphorylation sites on 78 proteins from *Bacillus subtilis*, which almost doubled the number of known bacterial phosphoproteins [26]. Further studies of this type are presently going on for *E. coli* and *Lactococcus lactis* (Macek et al., unpublished results), and taken together they promise to reshape our view of bacterial protein phosphorylation on serine/threonine/tyrosine residues.

2. Advantages of site-specific and gel-free approach in bacterial phosphoproteomics

The traditional phosphoproteomic approach with 2D gels followed by mass spectrometry peptide identification is still widely used today, and it has provided many interesting insights over the past years [8,27–29]. The main strength of this approach lies in its capacity to separate the complex protein mixtures, and one of its main limitations was the performance of mass spectrometry machines used to handle such samples. The nature of labelling (radiolabelling or detection with specific anti-phosphoamino acid antibodies) and the dynamic range and accuracy of mass spectrometry detectors usually did not suffice to identify non-abundant phosphoproteins and phosphoproteins with low occupancy of phosphorylation sites (such as bacterial proteins phosphorylated on tyrosine). However, with automated coupling of liquid chromatography systems to high-resolution mass spectrometers, 2D-gel separation is no longer a prerequisite for analyzing complex protein mixtures [31,32]. In this new approach, the crude protein extract is first digested to peptides (with trypsin or LysC) and subsequently enriched for phosphopeptides. The peptide mixture is thereafter separated by liquid chromatography, and the eluted fractions are directly analyzed with an Ion trap/Orbitrap hybrid mass spectrometer. This approach allowed the rapid and efficient detection of phosphopeptides from complex mixtures in a single experiment, also including tyrosine phosphorylated peptides from bacterial samples [31].

The phosphoproteome is expected to be dynamic, responding to different external and internal stimuli [28]. Therefore, analyzing how the phosphoproteome changes in different growth conditions or under stress is expected to provide a deeper insight into the mechanisms and functionality of protein phosphorylation. A technique known as stable isotope labeling by amino acids in cell culture (SILAC) is the latest development in the field of quantitative proteomics/phosphoproteomics [33]. It necessitates growing two populations of cells under different growth conditions or exposed to different stimuli. One population is grown in the presence of normal (light) amino acids, while the other is grown with heavy amino acids [34], containing ^{13}C, ^{15}N, or ^{2}H. After the proteins are harvested from the two-cell populations, mixed and digested, their peptides can be distinguished from each other as the heavy amino acids incorporated in the proteins compared to the light will be seen as mass shift in the spectra upon MS analysis [34]. The relative intensities of light and heavy peaks for the phosphorylated and nonphosphorylated form of a certain peptide allow the comparison of the extent of phosphorylation in the two conditions. Therefore, SILAC combined with high throughput and high resolution mass spectrometry will allow not only the site-specific detection of phosphorylation events, but also quantification of the occupancy of phosphorylation sites in bacterial cultures in response to various stimuli.

3. Distribution of phosphorylation sites: functional insights

Physiological significance of phosphorylation of bacterial proteins is versatile. Phosphorylation can activate or inhibit certain enzymes [3,12,13], target proteins for degradation [29], or affect their interactions with other cellular components [35]. In each of those cases, it can be argued that protein phosphorylation has a regulatory role, and it is therefore largely

acknowledged that serine/threonine/tyrosine phosphorylation plays an important role in the bacterial cell [36]. However, some sites might be phosphorylated non-specifically, by kinases with relaxed specificity, and have no regulatory role. In Fig. 1, the functional distribution of phosphorylation sites in *B. subtilis* reveals that majority of phosphorylated proteins have an enzymatic role. Functional assignment was performed manually, from primary scientific publications. Interestingly, proteins encoded by essential genes have been reported as over-represented among the phosphoproteins [31]. In the phosphorylated protein dataset, approximately one quarter of enzymes are involved in pathways of carbon metabolism. Other prominent groups include enzymes involved in DNA and protein metabolism and stress response. In addition, the *B. subtilis* phosphoproteome also contains a significant portion of phosphorylated proteins with unknown functions.

A number of key regulatory proteins in the transition growth phase of *B. subtilis* were found to be phosphorylated. When a growing *B. subtilis* culture slows down from exponential growth due to lack of nutrients it triggers a cascade of survival mechanisms [37]. Two very important players in this complex regulation, the global transcription regulator CodY [38] and a two-component system sensory kinase DegS [39] were found to be phosphorylated on serine residues. In case of CodY, the phosphorylated serine 215 is immediately flanking the DNA-binding motif, and it is very tempting to assume that its phosphorylation might repel the negatively charged DNA backbone. In case of DegS, which is implicated in competence regulation and salt-related stress response, the question is whether this phosphorylation is auto-inflicted or not, and whether it might affect the canonic phosphotransfer between DegS histidine and DegU aspartate. Another interesting perspective emerged concerning the key cellular regulator HPr. This small protein was known to be phosphorylated on histidine 15 and serine 46. Histidine phosphorylation participates in the phosphotransfer via the PTS, whereas serine phosphorylation triggers HPr to interact with the transcription regulator CcpA and elicit carbon catabolite regulation [40]. In the *B. subtilis* phosphoproteome, HPr phosphorylation on serine 46 was confirmed, but serine 12 was also found to be phosphorylated. This secondary phosphorylation might be incompatible with phosphorylation of histidine 15 since they are in immediate vicinity, and it might therefore separate the PTS role of HPr from its other regulatory roles.

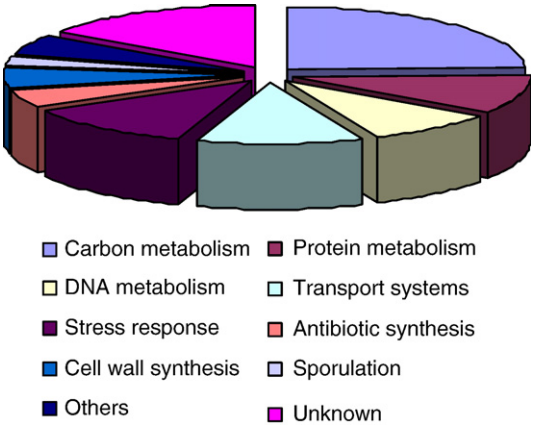


Fig. 1. Functional distribution of phosphoproteins in the *Bacillus subtilis* phosphoproteome [31]. Carbon metabolism 19 (9 in glycolysis, 4 in TCA cycle, 2 in the pentose phosphate pathway, 4 in other pathways), protein metabolism 8 (amino acid biosynthesis 3, protein synthesis 5), DNA metabolism 7, transport systems 9 (PTS 5, other systems 4), stress response 9 (sigma B regulators 6, others 3), antibiotic synthesis 4, cell wall synthesis 4, sporulation 2, other functions 4, unknown functions 12.

As outlined in Fig. 2, the subset of phosphoenzymes involved in carbon metabolism seems to be "strategically" distributed throughout various pathways, including both central pathways such as glycolysis and the TCA cycle, and also sugar import and exopolysaccharide biosynthesis. Could these phosphorylations have a regulatory role? If so, they may act as tap-like devices to redirect metabolic fluxes in accordance to different growth conditions. Moreover, central pathways of carbon metabolism have been proposed to have a more profound regulatory role than just processing of metabolites for energy extraction. For example, glycolysis has been shown to modulate some aspects of DNA replication, possibly signalling the energetic state of the cell [41]. Protein phosphorylation has the potential to add a whole new layer of regulation to cellular processes, but to which extent this is true in bacteria still needs to be ascertained experimentally. A very feasible approach of screening for phosphorylations that are "meaningful" from the regulatory aspect would include the use of SILAC in different growth conditions. Phosphorylation sites that change their occupancy significantly upon a change in growth conditions may be considered as "responsive", and thereby likely to regulate some sort of cellular response to adapt to the change.

4. Linking protein kinases and their substrates: the big challenge ahead

Bacteria posses a versatile repertoire of protein kinases capable of phosphorylating proteins on serine/threonine/tyrosine residues [36]. These comprise homologues of eukaryal kinases with so-called Hanks motifs [42], and a number of idiosyncratic bacterial kinases. Bacterial kinase "types" include bifunctional kinases, such as IDH kinase/phosphatase [3] and HPr kinase/phosphorylase [5], the BY-kinases [11], some kinases more or less related to sensory kinases from two-component systems, but capable of phosphorylating serine/threonine/tyrosine residues [43,44], and finally the tyrosine kinases with the arginino-phosphotransferase domain [45]. Many of these different kinases have been shown in individual biochemical studies to be specific for one specific substrate [26]. Nevertheless, is it possible to generalize this view, and assume that each kinase is specific for just one substrate? For example, it was shown in yeast that each kinase of the serine/threonine/tyrosine type phosphorylates on average 15 different substrates [46]. In *B. subtilis*, the characterized kinase complement is also largely outnumbered by the number of phosphorylated proteins [31]. Therefore, one has to assume that either many of these kinases are promiscuous, phosphorylating

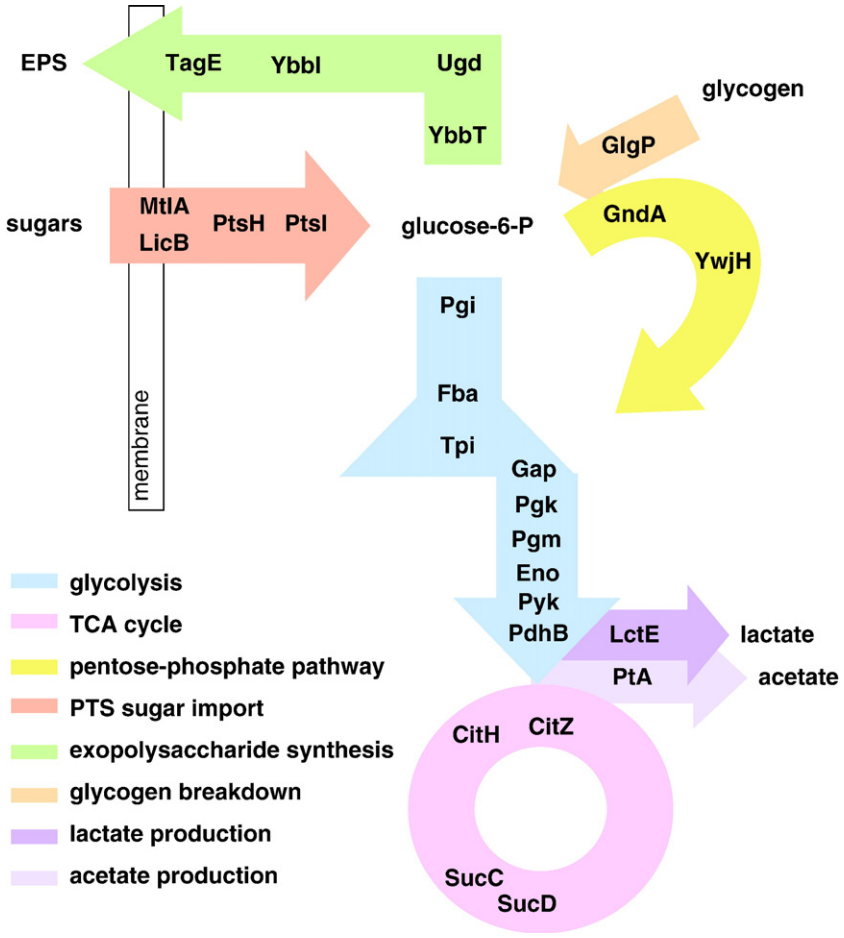


Fig. 2. Among phosphoproteins detected in *Bacillus subtilis*, there is a significant over-representation of enzymes involved in carbon metabolism [31]. Distribution of these phospho-enzymes among different biochemical pathways is represented schematically, the name of the phosphorylated enzyme indicated its placement in a metabolic pathway.

more than one substrate, such as for example the BY-kinase PtkA from *B. subtilis* [13,14], or there may be many yet uncharacterized kinases in the bacterial cell.

5. Bacteria-specific phosphorylation predictor?

The inventory of databases with phosphorylated protein residues grows steadily due to considerable experimental efforts in recent years. This naturally facilitated the development of bioinformatics in the field of phosphoproteomics. One of the principal challenges therein lies in developing reliable predictors of protein phosphorylation. Several different approaches have been employed, leading to different predictors available on the World Wide Web. The GPS predictor relies on group-based phosphorylation scoring, grouping the phosphorylation sites recognised by 71 different classes of protein kinases, and assigning the query protein to a specific group based on sequence homology [47]. DISPHOS (DISorder-enhanced PHOSphorylation predictor) exploits the finding that phosphorylated residues are often situated in intrinsically disordered protein regions to improve the assignment of phosphorylation sites [48]. The most widely used phosphorylation predictor, NetPhos, is based on neural networks alogorithms trained mainly on mammalian phosphorylation datasets [49]. The dataset employed in training of neural network algorithms naturally carries a bias towards a specific phyletic group. The bias of the original NetPhos proved to be significant enough to solicit the creation of NetPhosYeast [50] that performed much more accurately with respect to yeast protein phosphorylation. Bacteria, with their idiosyncratic kinases [5,11,41,51], present an even greater challenge for NetPhos and other homology-based predictors trained on eukaryal datasets. There are instances where NetPhos predictions proved of great value in bacterial systems [16], however, its general applicability on bacterial systems is largely questioned. Here we performed NetPhos predictions on all proteins in the recently published *B. subtilis* phosphoproteome [31], containing over 100 phosphorylated peptides (Table 1). Interestingly, only around 14% of all experimentally verified phosphorylation sites were recognised as the top scoring sites in their respective proteins. Furthermore, 50.5% of all sites scored a factor under 0.5 in NetPhos predictions, that would predict them as non-phosphorylated. It seems therefore that the capacity of NetPhos to predict phosphorylation sites in

Table 1
NetPhos predictions were performed on all proteins in the published *B. subtilis* phosphoproteome [31]

Protein	Phosphorylated residue	NPS	Protein	Phosphorylated residue	NPS	Protein	Phosphorylated residue	NPS
AhpF	*S58/59*	**0.82**	Ndk	S123	**0.99**	SucC	S220	0.00
AhpF	*Y393/S394*	0.17	Ndk	T92	0.02	SucD	*S19/T20*	0.01
AroA	S2	0.03	Ndk	*S118/120*	**1.00**	TagE	S2	0.00
Asd	S98	0.24	OppA	T470	0.11	TpiA	S213	0.36
Asd	Y146	0.02	OppA	*Y301/303*	0.02	Tsf	S149	0.31
CitH	S149	0.00	PdhB	*S302/306*	**0.99**	Ugd	*S69/Y70*	**0.96**
CitZ	S284	0.01	Pgi	T39	0.21	YbbI	S2	0.00
CodY	S215	0.01	Pgk	S183	**0.70**	YbbT	S100	0.01
CysS	S270	**0.69**	Pgk	T239	0.10	YcnE	S24	0.17
DegS	S76	0.43	Pgm	S62	0.31	YerA	S399	**0.60**
DhbC	S271	**0.70**	PnbA	S189	**0.61**	YerB	*T97/S103*	**0.87**
DhbF	S996	0.18	PrkC	T290	**0.55**	YfiY	S290	0.42
Drm	*T87/89*	0.20	PunA	S28	0.01	YfkK	S57	**0.53**
Eno	T141	0.02	Pta	*S128/T129*	**1.00**	YfnI	*T297/S298*	**0.98**
Eno	S259	0.15	PtsH	S12	0.05	YjdD	S365	**0.99**
Eno	S325	0.03	PtsH	S46	**0.97**	YjoA	Y150	0.10
Eno	*Y281/S285*	**1.00**	PtsI	S34	**0.99**	YnfE	Y12	**0.64**
FbaA	T212	0.12	Pyk	S36	**0.99**	YojH	S174	0.14
FbaA	T234	0.03	PyrB	S303	0.00	YojH	T186	**0.85**
FusA	S569	**0.98**	RocA	*T2/4/Y5*	0.13	YorK	Y473	**0.73**
FusA	S213	**1.00**	RsbR	*T168/171*	**0.96**	YpfD	S243	**0.85**
FusA	T43	0.23	RsbR	*T205/S215*	0.17	YpoC	S117	**0.91**
FusA	S679	**0.99**	RsbS	S59	0.01	YpsB	*S73/T75*	0.14
FusA	*T24/25*	**0.66**	RsbV	S56	0.01	YqbO	*S970/972*	**1.00**
FusA	S302	**1.00**	RsbV	S57	0.35	YqfN	S48	**0.54**
GapA	*S148/151/T153/154*	0.15	RsbV	S52	**0.98**	YqhA	S160	0.04
GlgP	*T291/S294*	**0.81**	SodA	T70	0.05	YqhA	T181	**0.94**
GndA	*T78/80*	0.42	SodA	T34	0.08	YtnP	S11	**0.72**
Hbs	T4	0.01	SpoIIAA	S58	0.04	YtxJ	*T93/99/S94/96*	0.11
InfA	Y60	**0.99**	spoVG	*S66/67*	**0.92**	YvyG	Y49	**0.93**
IspU	T186	**0.89**	SrfAA	S1006	**0.97**	YwfI	*S35/36/41/T47*	**0.99**
Ldh	Y224	**0.58**	SrfAB	S999	**0.93**	YwjH	S39	**0.95**
LicB	S37	0.06	SrfAB	S2045	**0.91**	YxxG	*Y102/S105*	**0.87**
MtlA	S559	**0.77**	SrfAC	S1003	**0.89**			

For each experimentally determined phosphorylation site, NetPhos score (NPS) is given. Scores under 0.5 are in normal typeface, above 0.5 in bold typeface and if they represent the top score in the entire protein they are bold and underlined. Scores of 1.00 resulted from rounding up of the fourth decimal (from 0.995 to 0.999), no actual 1.000 scores were obtained. In cases where the phosphorylation could not be pin-pointed to a single residue, (ambiguity from the MS data, residues italicized), the NetPhos prediction for the highest scoring one among the potential phospho-residues is shown.

B. subtilis is hardly better than random. This clearly advocates the need for a bacteria-specific version of such a predictor that should be trained on data from the emerging bacterial phosphoproteomes obtained with site-specific techniques.

6. Conclusions and perspectives

Current advances in phosphoproteomics [31–33] brought about for the first time the possibility to study bacterial phosphoproteomes in a site-specific manner. At this point there is still only limited data available [31], but site-specific bacterial phosphoproteomics is picking up momentum. The major conclusions from the study on *B. subtilis* [31], concerning the number and functional distribution of the phosphorylated proteins, seem to be confirmed also in *E. coli* and *L. lactis* (Macek et al., unpublished results). After the initial charting of phosphoproteomes of some key bacterial models, it will be interesting to see how these phosphoproteomes respond to changes in environmental conditions, by utilizing quantitative phosphoproteomics (SILAC). Such data will feed directly into bioinformatic models, not only the ones designed to predict protein phosphorylation sites, but also those aiming to model the global regulatory networks relying on protein phosphorylation. The abundance of new and site-specific data in bacterial phosphoproteomics is expected to enable us to tackle some global questions about bacterial protein phosphorylation at the level of systems biology. Modelling of signal transduction cascades in bacteria was previously performed for two-component systems [52]. In order for this kind of analysis to become feasible with bacterial serine/threonine/tyrosine phosphorylation, further work is needed to assign the new-found phosphorylated proteins to their cognate kinases and phosphatases and quantify the phosphorylation events.

Acknowledgments

This work was supported by the Danish Natural Science Research Council (FNU).

References

[1] T. Pawson, J.D. Scott, Protein phosphorylation in signaling–50 years and counting, Trends Biochem. Sci. 30 (2005) 286–290.

[2] J.A. Hoch, Two-component and phosphorelay signal transduction, Curr. Opin. Microbiol. 3 (2000) 165–170.

[3] M. Garnak, H.C. Reeves, Phosphorylation of isocitrate dehydrogenase of *Escherichia coli*, Science 203 (1979) 1111–1112.

[4] D.C. LaPorte, T. Chung, A single gene codes for the kinase and phosphatase which regulate isocitrate dehydrogenase, J. Biol. Chem. 260 (1985) 15291–15297.

[5] J. Deutscher, M.H. Saier, ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in *Steptococcus pyogenes*, Proc. Natl. Acad. Sci. U. S. A. 80 (1983) 6790–6794.

[6] J. Reizer, M.J. Novotny, C. Panos, M.H. Saier, Mechanismof inducer explulsion in *Streptococcus pyogenes*: a two-step processivated by ATP, J. Bacteriol. 156 (1983) 354–361.

[7] M. Kravanja, R. engelmann, V. Dossonnet, M. Blüggel, H.E. Meyer, R. Frank, A. Galinier, J. Deutscher, N. Schnell, W. Hengstenberg, The *hprK* gene of *Enterococcus faecalis* encodes a novel bifunctional enzyme: the HPr kinase /phosphatase, Mol. Microbiol. 31 (1999) 59–66.

[8] I. Mijakovic, S. Poncet, A. Galinier, V. Monedreo, S. Fieulaine, J. Janin, S. Nessler, J.A. Marquez, K. Scheffzek, S. Hasenbien, W. Hengstenberg, J. Deutcher, Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: a relic of early life? Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 13442–13447.

[9] J.C. Cortay, C. Rieul, B. Duclos, A.J. Cozzone, Characterization of the phosphoproteins of *Escherichia coli* cells by electrophoretic analyzes, Eur. J. Biochem. 159 (1986) 227–237.

[10] B. Duclos, C. Grangeasse, E. Vaganay, M. Riberty, A.J. Cozzone, Autophosphorylation of a bacterial protein at tyrosine, J. Mol. Biol. 259 (1996) 891–895.

[11] C. Grangeasse, A.J. Cozzone, J. Detscher, I. Mijakovic, Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology, Trends Biochem. Sci. 32 (2007) 86–94.

[12] G. Klein, C. Dartigalongue, S. Raina, Phosphorylation-mediated regulation of heat shock response in *Escherichia coli*, Mol. Microbiol. 48 (2003) 269–285.

[13] I. Mijakovic, S. Poncet, G. Boel, A. Maze, S. Gillet, E. Jamet, P. Decottignies, C. Grangeasse, P. Doublet, P. Le Marechal, J. Deutscher, Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases, EMBO J. 22 (2003) 4709–4718.

[14] I. Mijakovic, D. Petranovic, B. Macek, T. Cepo, M. Mann, J. Davies, P.R. Jensen, D. Vujaklija, Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine, Nucleic Acids Res. 34 (2006) 1588–1596.

[15] S. Lacour, P. Doublet, B. Obadia, A.J. Cozzone, C. Grangeasse, A novel role for protein-tyrosine kinase Etk from *Escherichia coli* K-12 related to polymyxin resistance, Res. Microbiol. 157 (2006) 637–641.

[16] Z. Minic, C. Marie, C. Delorme, J.M. Faurie, G. Mercier, D. Ehrlich, P. Renault, Control of EpsE, the phosphoglycosyltransferase initiating exopolysaccharide synthesis in *Streptococcus thermophilus*, by EpsD tyrosine kinase, J. Bacteriol. 189 (2007) 1351–1357.

[17] J. Munoz-Dorado, S. Inouye, M. Inouye, A gene encoding a protein serine/threonine kinase is required for normal development of *Myxococcus xanthus*, a gram-negative bacterium, Cell 67 (1991) 995–1006.

[18] E. Madec, A. Stensballe, S. Kjellstrom, L. Cladiere, M. Obuchowski, O.N. Jensen, S. Seror, Mass spectrometry and site-directed mutagenesis identify several autophosphorylated residues required for the activity of PrkC, a Ser/Thr kinase from *Bacillus subtilis*, J. Mol. Biol. 330 (2003) 459–472.

[19] V. Molle, D. Soulat, J.M. Jault, C. Grangeasse, A.J. Cozzone, J.F. Prost, Two FHA domains on an ABC transporter, Rv1747, mediate its phosphorylation by PknF, a Ser/Thr protein kinase from *Mycobacterium tuberculosis*, FEMS Microbiol. Lett. 234 (2004) 215–223.

[20] S. Backert, M. Selbach, Tyrosine-phosphorylated bacterial effector proteins: the enemies within, Trends Microbiol. 13 (2005) 476–484.

[21] J.L. Spudich, W. Stoeckenius, Light-regulated retinal-dependent reversible phosphorylation of Halobacterium proteins, J. Biol. Chem. 255 (1980) 5501–5503.

[22] R. Skorko, Protein phosphorylation in the archaeabacterium *Sulfolobus acidoaldarius*, Eur. J. Biochem. 145 (1984) 617–622.

[23] S.J. Jeon, S. Fujiwara, M. Takagi, T. Tanaka, T. Imanaka, Tk-PTP protein tyrosine/serine phosphatase from hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1: enzymatic characteristics and identification of its substrate protein, Biochem. Biophys. Res. Commun. 295 (2002) 508–514.

[24] P.J. Kennelly, Archaeal protein kinases and protein phosphatases: insights from genomics and biochemistry, Biochem. J. 370 (2003) 373–389.

[25] J. Eichleri, M.W.W. Adams, Posttranslational protein modification in Archaea, Mirco. Mol. Biol. Rev. 69 (2005) 393–425.

[26] S.M. Wurgler-Murphy, D.M. King, P.J. Kennelly, The phosphorylation site database: a guide to the serine-, threonine-, and/or tyrosine-phosphorylated proteins in prokaryotic organisms, Proteomics 4 (2004) 1562–1570.

[27] A.K. Bendt, A. Burkovski, S. Schaffer, M. Bott, M. Farwick, T. Hermann, Towards a phosphoproteome map of *Corynebacterium glutamicum*, Proteomics 3 (2003) 1637–1646.

[28] A. Levine, F. Vannier, C. Absalon, L. Kuhn, P. Jackson, E. Scrivener, V. Labas, J. Vinh, P. Courtney, J. Garin, S.J. Seror, Analysis of the dynamic *Bacillus subtilis* Ser/Thr/Tyr phosphoproteome implicated in a wide variety of cellular processes, Proteomics 6 (2006) 2157–2173.

[29] R. Rosen, D. Becher, K. Buttner, D. Biran, M. Hecker, E.Z. Ron, Highly phosphorylated bacterial proteins, Proteomics 4 (2004) 3068–3077.

[30] Q. Hu, R. Noll, L. Hongyan, A. Makarov, M. Hardman, G. Cooks, The Orbitrap: a new mass spectrometer, J. Mass Spectrom. 40 (2005) 430–443.

[31] B. Macek, I. Mijakovic, J.V. Olsen, F. Gnad, C. Kumar, P.R. Jensen, M. Mann, The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*, Mol. Cell. Proteomics 6 (2007) 697–707.

[32] J. Olsen, B. Blagoy, F. Gnad, B. Macek, C. Kumar, P. Mortensen, M. Mann, M. Global, In vivo, and site-specific phosphorylation dynamics in signalling networks, Cell 127 (2006) 635–648.

[33] S. Ong, B. Blagoev, I. Kratchmarova, D. Kristensen, H. Steen, A. Pandey, M. Mann, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics, Mol. Cell. Proteomics 1 (2002) 376–386.

[34] M. Mann, Functional and quantitative proteomics using SILAC, Nature 7 (2006) 952–958.

[35] C. Alexander, N. Bilgin, C. Lindschau, J.R. Mesters, B. Kraal, R. Hilgenfeld, V.A. Erdmann, C. Lippmann, Phosphorylation of elongation factor Tu prevents ternary complex formation, J. Biol. Chem. 270 (1995) 14541–14547.

[36] J. Deutscher, M.H. Saier, Ser/Thr/Tyr protein phosphorylation in bacteria – for long time neglected, now well established, J. Mol. Microbiol. Biotechnol. 9 (2005) 125–131.

[37] M.A. Strauch, J.A. Hoch, Transition-state regulators: sentinels of *Bacillus subtilis* post-exponential gene expression, Mol. Microbiol. 7 (1993) 337–342.

[38] V.M. Levdikov, E. Blagova, P. Joseph, A.L. Sonenshein, A.J. Wilkinson, The structure of CodY, a GTP- and isoleucine-responsive regulator of stationary phase and virulence in gram-positive bacteria, J. Biol. Chem. 281 (2006) 11366–11373.

[39] F. Kunst, T. Msadek, J. Bignon, G. Rapoport, The DegS/DegU and ComP/ComA two-component systems are part of a network controlling degradative enzyme synthesis and competence in *Bacillus subtilis*, Res. Microbiol. 145 (1994) 393–402.

[40] J. Deutscher, C. Francke, P.W. Postma, How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria, Microbiol. Mol. Biol. Rev. 70 (2006) 939–1031.

[41] L. Janniere, D. Canceill, C. Suski, S. Kanga, B. Dalmais, R. Lestini, A.F. Monnier, J. Chapuis, A. Bolotin, M. Titok, E.L. Chatelier, S.D. Ehrlich, Genetic evidence for a link between glycolysis and DNA replication, PLoS ONE 2 (2007) e447.

[42] S.K. Hanks, A.M. Quinn, T. Hunter, The protein kinase family: conserved features and deduced phylogeny of the catalytic domains, Science 241 (1988) 42–52.

[43] J. Wu, N. Ohta, J.L. Zhao, A. Newton, A novel bacterial tyrosine kinase

essential for cell division and differentiation, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 13068–13073.
[44] S.M. Najafi, A.C. Willis, M.D. Yudkin, Site of phosphorylation of SpoIIAA, the anti-antisigma factor for sporulation-specific sigma F of *Bacillus subtilis*, J. Bacteriol. 177 (1995) 2912–2913.
[45] J. Kirstein, D. Zuhlke, U. Gerth, K. Turgay, M. Hecker, A tyrosine kinase and its activator control the activity of the CtsR heat shock repressor in B. subtilis, EMBO J. 24 (2005) 3435–3445.
[46] J. Ptacek, G. Devgan, G. Michaud, H. Zhu, X. Zhu, J. Fasolo, H. Guo, G. Jona, A. Breitkreutz, R. Sopko, R.R. McCartney, M.C. Schmidt, N. Rachidi, S.J. Lee, A.S. Mah, L. Meng, M.J. Stark, D.F. Stern, C. De Virgilio, M. Tyers, B. Andrews, M. Gerstein, B. Schweitzer, P.F. Predki, M. Snyder, Global analysis of protein phosphorylation in yeast, Nature 438 (2005) 679–684.
[47] Y. Xue, F. Zhou, M. Zhu, K. Ahmed, G. Chen, X. Yao, GPS: a comprehensive www server for phosphorylation sites prediction, Nucleic Acids Res. 33 (2005) W184–W187 Web Server issue.
[48] L.M. Iakoucheva, P. Radivojac, C.J. Brown, T.R. O'Connor, J.G. Sikes, Z. Obradovic, A.K. Dunker, The importance of intrinsic disorder for protein phosphorylation, Nucleic Acids Res. 32 (2004) 1037–1049.
[49] N. Blom, S. Gammeltoft, S. Brunak, Sequence and structure-based prediction of eukaryotic protein phosphorylation sites, J. Mol. Biol. 294 (1999) 1351–1362.
[50] C.R. Ingrell, M.L. Miller, O.N. Jensen, N. Blom, NetPhosYeast: prediction of protein phosphorylation sites in yeast, Bioinformatics 23 (2007) 895–897.
[51] C. Oudot, J.C. Cortay, C. Blanchet, D.C. Laporte, A. Di Pietro, A.J. Cozzone, J.M. Jault, The "catalytic" triad of isocitrate dehydrogenase kinase/phosphatase from *E. coli* and its relationship with that found in eukaryotic protein kinases, Biochemistry 40 (2001) 3047–3055.
[52] M. Kollmann, L. Løvdok, K. Bartholome, J. Timmer, V. Sourjik, Design principles of a bacterial signalling network, Nature 438 (2005) 504–507.

EXPERT REVIEWS

Phosphoproteomics in bacteria: towards a systemic understanding of bacterial phosphorylation networks

Expert Rev. Proteomics 5(4), 619–627 (2008)

Carsten Jers, Boumediene Soufi, Christophe Grangeasse, Josef Deutscher and Ivan Mijakovic†

†*Author for correspondence*
Technical University of Denmark, Biosys, Center for Microbial Biotechnology, DK-2800 Lyngby, Denmark
Tel.: +45 4525 2495
Fax: +45 4593 2809
im@bio.dtu.dk

Bacteria use protein phosphorylation to regulate all kinds of physiological processes. Protein phosphorylation plays a role in several key steps of the infection process of bacterial pathogens, such as adhesion to the host, triggering and regulation of pathogenic functions as well as biochemical warfare; scrambling the host signaling cascades and impairing its defense mechanisms. Recent phosphoproteomic studies indicate that the bacterial protein phosphorylation networks could be more complex than initially expected, comprising promiscuous kinases that regulate several distinct cellular functions by phosphorylating different protein substrates. Recent advances in protein labeling with stable isotopes in the field of quantitative mass spectrometry phosphoproteomics will enable us to chart the global phosphorylation networks and to understand the implication of protein phosphorylation in cellular regulation on the systems scale. For the study of bacterial pathogens, in particular, this research avenue will enable us to dissect phosphorylation-related events during different stages of infection and stimulate our efforts to find inhibitors for key kinases and phosphatases implicated therein.

KEYWORDS: bacterial pathogen • cellular adhesion • gene regulation • kinase • mass spectrometry • phosphatase • phosphoproteomics • protein phosphorylation • signaling • virulence

Protein phosphorylation is globally recognized as an important regulatory and signaling mechanism. Historically, it was first described in *Eukarya* [1], but during the last two decades it has been clearly established that this post-translational modification is ubiquitous [2]. The most common types of protein phosphorylation (namely on serine, threonine, tyrosine, histidine and aspartate residues) have been identified in all three kingdoms of life. In particular, histidine and aspartate phosphorylation are mainly confined to the signal-transduction couples known as the two-component systems [3], which will be only briefly treated here. Histidine phosphorylation is also prominent in bacterial phosphoenolpyruvate-dependent phosphotransferase systems, which have also been linked to bacterial pathogenicity. For example, several pathogens use mechanisms controlling carbohydrate utilization also for the regulation of virulence gene expression [4,5]. However, the main focus of this review will be on serine, threonine and tyrosine phosphorylation. Implication of serine/threonine and tyrosine kinases in the control of pathogenicity functions has been extensively documented in numerous bacterial pathogens (for reviews see [6–8]). Bacteria possess a versatile repertoire of protein kinases and phosphatases, comprising both ubiquitous enzymes, such as Hanks-type kinases [9] or PPP and PPM families of protein-serine/threonine phosphatases [10] and some idiosyncratic bacteria-specific enzymes, such as BY-kinases [11] or PHP-like phosphotyrosine phosphatases [12]. Many case studies with various pathogens have provided evidence of the role of these kinases and phosphatases in different stages of infection, ranging from attachment to host cells [13,14], and involvement in gene regulation of pathogenic functions [15,16], to controlling [17] and actively participating in [18–20] the biochemical warfare against the host signaling and immune system (FIGURE 1). These case studies usually describe, in quite some detail, isolated control switches where a kinase exerts its physiological effects based on either autophosphorylation or

 10.1586/14789450.5.4.619 ISSN 1478-9450

phosphorylation of a single bacterial protein substrate implicated in a specific function. However, recent site-specific phosphoproteomic studies in the bacterial model organisms *Bacillus subtilis* [21] and *Escherichia coli* [22] indicated the existence of over 100 phosphorylation sites in each bacterium. The first phosphoproteomic study in a pathogenic bacterium, *Campylobacter jejuni*, indicated 58 phosphorylation sites [23]. This number largely surpasses the number of characterized kinases in these bacteria and suggests the possibility that either many bacterial kinases are promiscuous (i.e., phosphorylate several different substrates), or there could be a large number of as yet uncharacterized kinases not exhibiting known protein kinase folds that would allow the prediction of their function from the genomic data. There are arguments to support both of these hypotheses. Promiscuous kinases have indeed been described in bacteria, they are capable of phosphorylating several classes of protein substrates and their knockouts lead to pleiotropic phenotypes [24–30]. Conversely, there are also cases where new classes of protein kinases were discovered. They include proteins for which kinase activity would have been impossible to predict [31]. Be it as it may, bacterial phosphorylation networks appear to be larger than originally expected and, clearly, systems-level studies are called for to understand the implications of protein phosphorylation in bacterial regulation and physiology. In particular, phosphoproteomic approaches with bacterial pathogens promise to provide key insights into the implication of kinases and phosphatases in the virulence process and the basis for choosing attractive targets for rational-design drug development.

Bacterial two-component systems involved in virulence

In canonic two-component systems, the first component of this signaling 'switch' is an input-sensing histidine kinase that autophosphorylates upon activation via a specific stimulus. Thereafter, the kinase transmits the phosphate to an aspartate residue of the second component, the response regulator, which subsequently triggers the signal response, usually by activating gene transcription. Several two-component systems can be integrated in a more complex signaling cascade, known as a phospho-relay [3]. Two-component systems and phospho-relays are extensively used in virulence-related signaling (for reviews see [32,33]). The spectrum of pathogenicity-related functions under the control of two-component systems is very broad. It ranges from toxin production and adhesion [34], quorum sensing [35], capsule synthesis and motility [36] to drug resistance [37]. As discussed by Beier and Gross [33], one key aspect of two-component system-dependent bacterial virulence regulation is still insufficiently clear: the nature of the input signals for the histidine autokinases. The inputs recognized so far include ions (Mg^{2+}, Ca^{2+}, Fe^{3+}), pH value, oxygen availability, temperature and osmolarity. Interestingly, the input can also be constituted by serine/threonine phosphorylation, as in the case of *Streptococcus agalactiae*, where two-component system signaling converges with the serine/threonine kinase activity of Stk1 [38]. Histidine kinases and response regulators of two-component systems have very few described homologues in *Eukarya* and none in humans, so they represent promising selective targets for drug development [39].

Serine/threonine/tyrosine phosphorylation controls the expression of virulence genes

A series of studies has established that serine/threonine/tyrosine phosphorylation can control bacterial virulence at the level of gene expression. In *Streptococcus pneumoniae*, the inactivation of a Hanks-type serine/threonine kinase, StkP, leads to a decrease in expression of approximately 50 genes and an increase in expression of about the same number of genes [15]. This global effect on transcription was attributed to the capacity of StkP to phosphorylate the α subunit of the *S. pneumoniae* RNA polymerase. In *Mycobacterium tuberculosis*, the genes responsible for resistance to the antimycobacterial agent ethambutol are controlled by the transcription activator EmbR, which is in turn controlled via phosphorylation by several Hanks-type serine/threonine kinases including PknH [16]. In addition to EmbR, other substrates of PknH (Rv0681 and DacB1) were suggested to participate in gene regulation of processes that affect the survival of *M. tuberculosis* in murine infection systems [26]. In the enteropathogenic *E. coli*, a novel member of the ribosome-binding GTPase superfamily, BipA, was found to autophosphorylate on tyrosine [40] and participate in diverse regulatory cascades to co-ordinate the expression of key pathogenicity islands and other virulence-associated factors [41]. In *S. agalactiae*, the serine/threonine kinase Stk1 positively regulates the expression of a gene for β-hemolysin/cytolysin required for resistance of group B *Streptococci* to human blood [38]. In the aforementioned cases, the kinase directly phosphorylates the transcription factors/gene regulators thereby affecting their activity. In other cases, virulence regulators are controlled by specific interactions with the phosphorylated or dephosphorylated form of other regulatory proteins. This seems to be the case for the Crp-like transcription activator, PrfA, which controls the expression of several virulence genes including *hly, plcA, plcB, actA, inlA* and *inlB,* and which was proposed to be fully active only in the presence of one or several phosphorylated phosphotransferase system components [42,43]. Finally, phosphorylation-dependent co-repressors can control the synthesis of virulence regulators, as was shown for Mga from *Streptococcus pyogenes* [4]. Serine-46 phosphorylated HPr (the central regulator of carbon metabolism in *Firmicutes*) interacts with the transcription regulator CcpA, and thus stimulates *mga* and ultimately virulence gene expression in response to the serine phosphorylation of HPr by HPr kinase/phosphorylase. It seems that these bacterial pathogens have coupled sugar availability (which triggers the kinase activity of HPr kinase/phosphorylase) to the control of pathogenic functions [44]. One of the best known bacterial systems relying on signal transduction by serine/threonine kinases/phosphatases is the control of the alternative σ factors by phosphorylation-dependent partner-switching modules, described in *B. subtilis* [45]. Two linear switch modules each comprise a phosphatase, a kinase-switch protein and

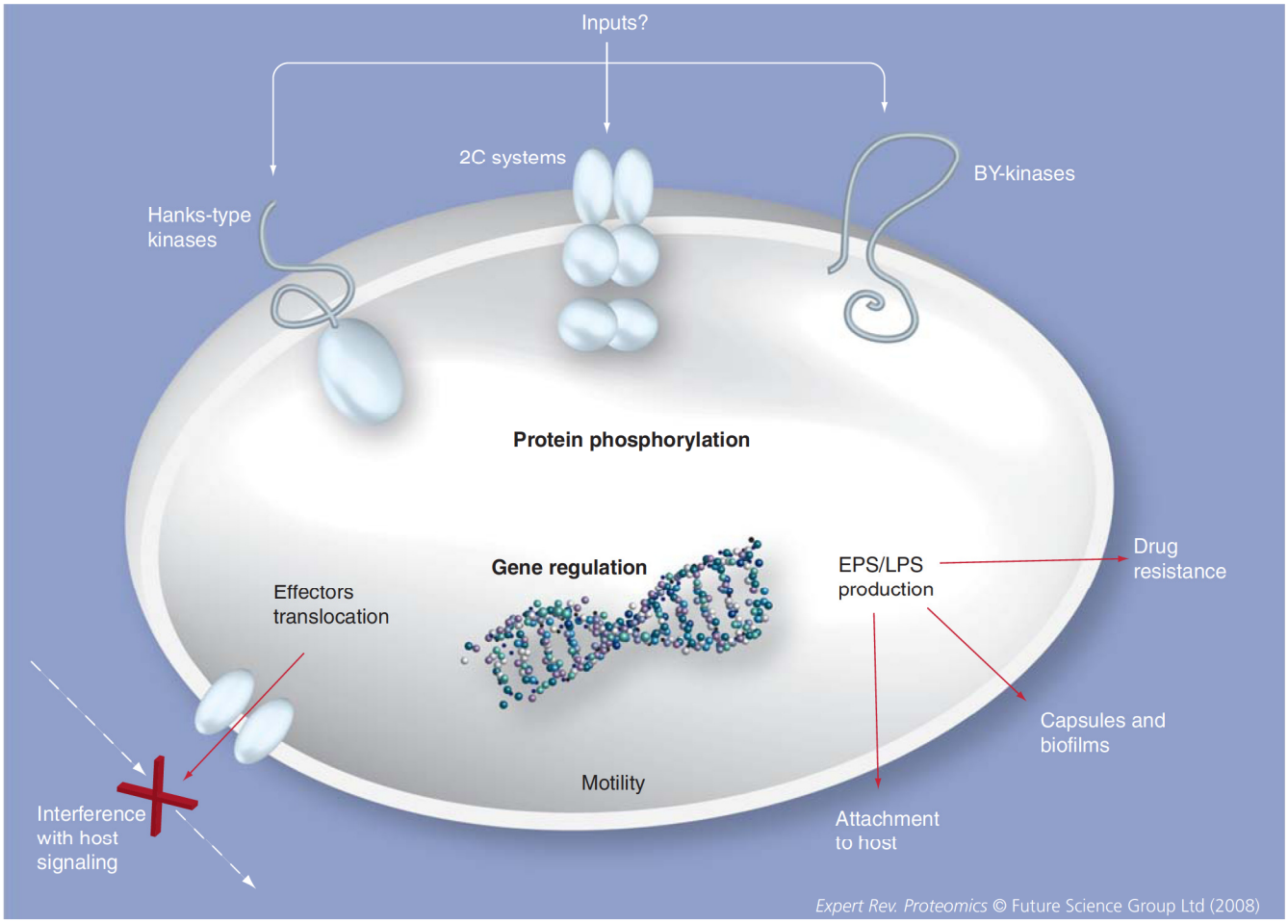


Figure 1. Overview of phosphorylation-related functions of bacterial pathogens. Upon infection, different types of bacterial kinases are triggered by presently poorly understood stimuli. These kinases unleash and mediate a variety of phosphorylation-dependent pathogenic functions. Those include transcriptional activation of pathogenicity factors, controlling the aspects of the cell surface through production of polysaccharides, capsules and biofilms, and finally biochemical warfare against the host, which implies export of virulence factors into the host cell where they scramble the host signaling cascades and the immune response.
EPS: Extracellular polysaccharide; LPS: Lipopolysaccharide.

an antagonist. They integrate various environmental signals and the ultimate kinase switch, RbsW, inhibits the σ^B factor by protein–protein interaction. σB has been found to control virulence functions in pathogenic bacteria such as *Listeria monocytogenes* [46,47], *Staphylococcus aureus* [48] and *Bacillus cereus* [49].

Serine/threonine/tyrosine phosphorylation controls the characteristics of the bacterial cell surface

The key to a successful invasion by bacterial pathogens lies in their ability to modify the physicochemical aspects of their cell surface; it is often at the cell surface that the battle with the host immune system is won or lost [50]. Pathogens modify their capsular and noncapsular exopolysaccharides in order to assure optimal attachment to the host cell, change the composition of their lipopolysaccharides to protect themselves from antibiotic agents, create dense extracellular matrices to form biofilms and expose decoy molecules on their surface to pacify the immune system. Many aspects of this behavior are actually controlled by bacterial protein phosphorylation. These regulatory phenomena can be divided into two classes. The first class can be grouped into events where transmembrane kinase autophosphorylation directly affects extracellular polysaccharide synthesis, assembly and translocation, usually via large transmembrane multiprotein complexes in which the kinase takes part [51]. Such a role is frequently played by BY-kinases in pathogenic and nonpathogenic bacteria alike [9,11,12,52,53]. In addition, it has been observed that the BY-kinases, such as Etk from *E. coli* [54], can be involved in the bacterial susceptibility to the cationic antimicrobial peptides by influencing lipopolysaccharide modifications. In the second class we can group all those events where kinases phosphorylate other proteins, usually enzymes catalyzing different steps in extracellular polysaccharide and lipopolysaccharide synthesis. In this role, we equally find the BY-kinases and the Hanks-type serine/threonine kinases.

For example, the *E. coli* BY-kinase, Wzc, is known to phosphorylate and thereby activate the enzyme Ugd, which produces the colanic acid precursor UDP-glucuronate [55]. A similar system has recently been described in *S. aureus*, where the BY-kinase Cap5B2 was found to activate the UDP-acetyl-mannosamine dehydrogenase Cap5O by tyrosine phosphorylation [56]. Hanks-type kinases also exhibit an analogous regulatory role. In *S. pneumoniae*, StkP was found to phosphorylate the phosphoglucosamine mutase, GlmM, which catalyzes the first step in the synthesis of the essential cell envelope component [57]. Three protein components of the type II fatty acid synthase from *M. tuberculosis*, namely mtFabD, KasA and KasB, were found to be phosphorylated by various serine/threonine kinases in this organism [58]. Curiously, KasA and KasB are regulated by phosphorylation in a differential manner, the former is inactivated and the latter is activated by phosphorylation. In some cases, even the biosynthetic precursors themselves get phosphorylated by protein kinases, such is the case of the unusual tyrosine kinase WaaP from *P. aeruginosa* that, in addition to autophosphorylation, also phosphorylates L-glycero-D-mannoheptose in the inner core region of lipopolysaccharides [59].

Serine/threonine/tyrosine phosphorylation interferes with host signaling

Pathogenic bacteria, especially intracellular pathogens, have developed a versatile arsenal for biochemical warfare against the host signaling cascades, allowing them to survive in this hostile environment [5,17,60]. This usually includes exporting kinases and phosphatases capable of interfering with the host phosphorylation-dependent signaling cascades. Already, the export of these 'weapons' from the bacterial cell is regulated by phosphorylation. For example, in *P. aeruginosa*, threonine phosphorylation of the autokinase PpkA controls the activity of a type IV secretion system via a forkhead-associated domain [15,61]. In *Chlamydophila pneumoniae*, an interesting and novel dual-specificity kinase, PknD, was found to autophosphorylate on serine and tyrosine and phosphorylate the forkhead-associated domain of a type III secretion system on threonine and tyrosine [62]. In *Bordetella bronchiseptica*, synthesis of a type III secretion system is controlled by the previously mentioned phosphorylation-dependent σ^B partner-switching mechanism [63]. Once excreted into the host cell, effectors from bacterial pathogens attack a wide array of targets involved in host signaling and immunity. A serine/threonine kinase, PknG, of *M. tuberculosis* is secreted into the host macrophage cytosol and contributes to blocking the maturation of mycobacteria-containing phagosomes to lysosomes [20]. During infection, the MAPK pathway in host cells is normally activated and *Salmonella* secretes a specific phosphatase, SptP, to reverse this activation [64]. Indeed, phosphatases often seem to be the weapon of choice secreted by bacterial pathogens to disrupt host immunity. *Shigella flexneri* secretes a phosphatase, OspF, that specifically targets the pathway responsible for phosphorylation of host histone H3, thus blocking the activation of a subset of NF-κB-responsive genes [18]. *M. tuberculosis* secrets a phosphatase, PtpA, that inhibits phagosome–lysosome fusion by dephosphorylating the key regulator VPS33B [65]. The plant pathogen *Pseudomonas syringae* uses a tyrosine phosphatase HopAO1 to suppress a number of resistance genes in its host *Arabidopsis thaliana* [66]. *Yersinia* species are particularly versatile in their attack strategies. They excrete a variety of effector proteins, some of which are actin-activated kinases (YopO) [67], phosphatases (YopH) [68] and even acetyltransferases (YopJ) that prevent host kinase activation by acetylating the phosphorylation sites [69]. A similar strategy is employed by *P. aeruginosa* ExoS, an ADP-ribosyltransferase that inhibits ezrin/radixin/moesin phosphorylation by ADP-ribosylating several arginine residues flanking the phosphorylation site [70]. *Helicobacter pylori* uses yet another strategy, it excretes a protein, CagA, that is tyrosine phosphorylated by the host and then induces the dephosphorylation of host proteins, rearrangements of the host cell actin cytoskeleton and cell scattering, by inhibiting the activity of c-Src [71]. Similar roles are played by the protein Tir from enteropathogenic *E. coli* [72] and AnkA from *Anaplasma phagocytophilum* [73]. Any given bacterial pathogen is likely to use several phosphorylation-dependent strategies in the invasion process, as exemplified by *M. tuberculosis* [8], to assure its ultimate success in overcoming the host defenses. Interestingly, when infected with phages, key physiological functions in bacteria themselves can be the target of protein phosphorylation-mediated attacks. From a historical point of view, it is interesting to note that among the first phosphoproteins detected in bacteria was the *E. coli* RNA polymerase β´ subunit [74]. Together with several other proteins of the *E. coli* translation and transcription machineries, this enzyme becomes phosphorylated by the protein kinase Gp0.7 encoded by one of the early genes of the bacteriophage T7. The aim of the coordinated phage-induced protein phosphorylation is to stop the synthesis of bacterial proteins in favor of phage proteins. A related protein kinase is found in numerous other bacteriophages infecting, for example, *Yersinia*, *Kluyvera* and *Klebsiella*. The site of phosphorylation in the *E. coli* RNA polymerase β´ subunit was recently found to be Thr-1068 [75].

Systems approach to studying serine/threonine/tyrosine phosphorylation in pathogenic bacteria

Protein phosphorylation is a dynamic regulatory event. Since the development of mass spectrometry (MS) techniques capable of ionizing proteins and peptides during the 1990s, became the method of choice for analyzing protein phosphorylation. The development of this technology is beyond the scope of this review; readers interested in an overview of various methods are referred to an excellent review by Gafken and Lampe [76]. Techniques for global-scale phosphoproteome analysis are also reviewed by Collins *et al.* [77]. Recently developed fragmentation techniques, such as electron capture dissociation, are now capable of fragmenting the peptide bonds while preserving phosphorylation on serine and threonine residues, which is a promising breakthrough in phosphoproteomic studies [78]. So far, MS proteomics has been successfully used only on phospho-serine, -threonine and -tyrosine

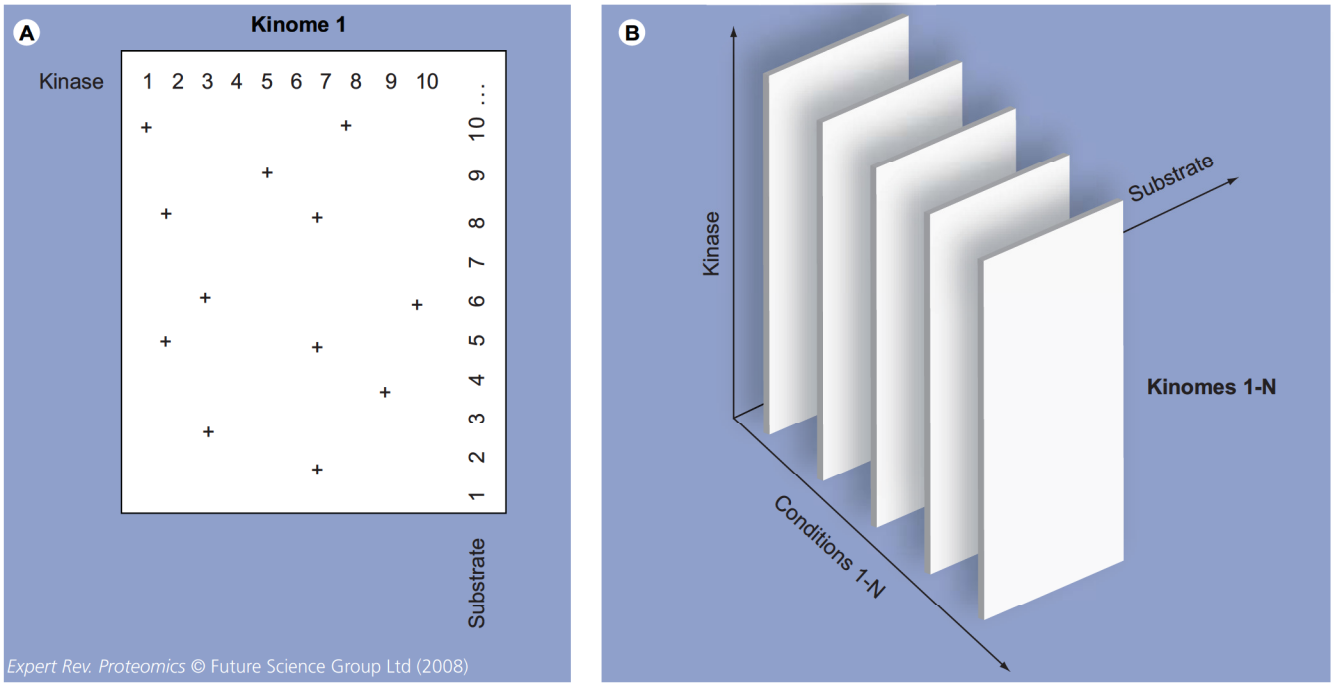


Figure 2. Charting the bacterial phosphorylation network using the quantitative mass spectrometry approach. (A) Creation of single kinase knockouts, coupled to quantitative mass spectrometry, will allow the tracing of links between individual kinases and their *in vivo* substrates. For any given condition, we can thus define a kinome; some kinases will have multiple substrates, and some will have none under employed experimental conditions. **(B)** Once we start altering the growth conditions, we can expect the phosphorylation patterns to change; some kinases will be shut down and others activated. Hence, for N different conditions we can expect N different kinomes. This type of systematic approach certainly implies a large number of time-consuming and costly experiments, but it promises crucial insights into bacterial physiology in return.

residues. Phospho-histidines and -aspartates, prominent players in bacterial two-component systems, are still elusive targets for global MS analyses. Their rapid dephosphorylation *in vivo* and chemical instability during phosphopeptide enrichment and MS analysis make them a challenging target for this type of analysis.

Bacteria in general and bacterial pathogens in particular often face a hostile environment during their life cycle, and their responses and adaptation mechanisms have to be rapid and efficient. Protein phosphorylation-related signaling is no exception to this rule, and the pattern of protein phosphorylation is expected to undergo significant changes over time. Dramatic recent advances in the technology of MS-based proteomics [79] made possible the first site-specific phosphoproteomic studies in bacteria [21,22]. These studies yielded long lists of phosphorylated proteins, but these are essentially only snapshots of a dynamic network and therefore, are only of limited value. Quantitative and time-resolved approaches are needed to truly understand phosphorylation-dependent signaling and regulation. Which kinases are active at what time and what are the substrates they phosphorylate? Recent technological advances in quantitative MS will enable us to tackle these questions. Quantitative proteomics requires the ability to compare certain characteristics or signals from a peptide exposed to different stimuli [80]. This requires the use of high-throughput MS techniques in combination with labeling proteins with stable isotopes [80]. Stable isotopic labeling of peptides has been used for decades and is the method of choice used in MS-based quantitative proteomics [81]. There are many ways to introduce a stable isotope to a protein sample. The first is referred to as absolute quantification (AQUA), which involves the chemical synthesis of the isotope that would serve as a standard as it is introduced to the cells in a known amount [80]. The limitation with AQUA lies in the fact that the internal standard used for labeling has to be synthesized specifically for each protein in a cell and therefore, is not practical in large-scale quantitative proteomics studies. Chemically tagging proteins is another method of introducing a stable isotope into a proteome [80]. Two common methods used in chemical tagging are ICAT [79] and the iTRAQ™ [80]. ICAT uses light and heavy forms of a chemical substance to label and quantify peptides but lacks the ability to completely label protein targets and is known to have unspecific interactions, which complicate identification and quantification by MS [82]. The iTRAQ method involves a chemical label containing a tag that will have a specific fragmentation pattern when subjected to MS analysis. Each protein sample will contain a different isobaric tag, and up to four different tags (corresponding to four different experimental conditions) are available. The only minor disadvantage of iTRAQ compared with metabolic labeling techniques lies in the fact that the labeling takes place after peptide isolation, which can be a source of some sample

loss and variation during parallel treatment [80]. SILAC is the newest and most powerful form of introducing a stable isotope into a cellular system [83,84]. In the case of quantifying post-translational modifications such as phosphorylation, SILAC can be used to measure the relative occupancy of phosphorylation sites in peptides isolated from two different strains, or a single strain grown under two different conditions. It is increasingly used for relative global-scale comparisons of proteomes and phosphoproteomes. The recent study of the insulin signaling pathway exemplifies the utility of this method in reconstructing a complex branched phosphorylation cascade, including many tyrosine kinases [85]. We have demonstrated that SILAC, can be successfully applied in bacterial systems [MACEK *ET AL.*. UNPUBLISHED DATA] and we expect this method to provide invaluable insights into the dynamic nature of protein phosphorylation in bacteria. For example, quantitative phosphoproteomes of strains with inactivated individual kinase genes will allow us to link kinases to their physiological substrates (FIGURE 2). Moreover, we will be able to follow the adaptations in the kinome in response to changing environmental conditions in the experimental setup. A further step ahead in the technology of quantitative phosphoproteomics will come with AQUA, for example by using elemental phosphorous standards [86].

Once the quantitative data become available, a sensible objective would be the modeling of the kinase–phosphatase-comprising networks. Such networks can be modeled using the existing tools developed for treating eukaryal data [87]. Systems biologists are increasingly aware that transcriptional regulation determines metabolic flux activities in only some cases and very often other layers of regulation come into play [88,89]. Post-translational control modules, such as a phosphorylation network, should be superimposed on metabolic networks in order to fully account for the system behavior *in vivo*. Phosphorylation network data should easily provide Boolean-type inputs for integration [88] and once the time-resolved phosphoproteomics data and data on the relationship between phosphorylation and enzyme activity become available, more complex kinetic models can be made. Solid models with real predictive power are rare, but a steady improvement in that field makes it a promising avenue in all aspects of biology, fighting bacterial pathogens included.

Expert commentary & five-year view

The gel-based phosphoproteomics of bacteria still provides useful information because one can follow the phosphorylation pattern of many cytosolic proteins (including kinetics), which is a simple and useful approach for kinetic or mutant studies. Recent advances in bacterial phosphoproteomics, namely high-accuracy MS, providing global and site-specific data, enabled us to treat bacterial protein phosphorylation at the systems level. With the advent of quantitative techniques, such as SILAC, we are now capable of monitoring the exact occupancy of hundreds of phosphorylation sites in different experimental setups. In bacterial pathogens, this will enable us to closely follow the implication of protein phosphorylation in the infection process and single out the kinases and phosphatases that trigger key pathogenic functions. In the coming years we expect to see quantitative phosphoproteomic analyses performed on bacterial pathogens, yielding high-quality data for modeling, and a deeper understanding of key steps in the invasion process.

Acknowledgements

We are particularly grateful to Boris Macek for insightful discussions regarding the comparison of quantitative methods in mass spectrometry phosphoproteomics.

Financial & competing interests disclosure

This work was supported by the grants from the Danish Natural Science Research Council (FNU) to Ivan Mijakovic and Agence Nationale de la Recherche (ANR-07-JCJC0125–01) to Christophe Grangeasse. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Key issues

- Protein phosphorylation on serine/threonine/tyrosine residues controls many aspects of bacterial pathogenicity, such as triggering the pathogenicity functions, regulating adhesion to host cells and attacking the host immune system.
- Idiosyncratic bacterial kinases (such as two-component histidine kinases and BY-kinases), having no eukaryal homologs, are promising targets for future inhibitors and antimicrobial drug development.
- It is quite probable that bacteria harbor as yet undiscovered classes of protein kinases that account for phosphorylating the large number of phosphoproteins revealed by recent phosphoproteomics studies.
- The bacterial phosphorylation-based signaling network is complex and likely to involve promiscuous kinases and phosphatases and, therefore, systems-based studies may be more promising than individual case studies.
- Systems-level approaches such as phosphoproteomics and network modeling are the next logical step towards the better understanding of bacterial pathogens.
- High-throughput, site-specific and quantitative phosphoproteomic studies on samples taken on a time scale that is relevant *in vivo* will provide key insights as to the involvement of protein phosphorylation in bacterial virulence.
- In the future, label-free quantitation approaches will probably replace labeling techniques in mass spectrometry phosphoproteomics.

References

Papers of special note have been highlighted as:

- • of interest
- •• of considerable interest

1 Fischer EH, Krebs EG. Conversion of phosphorylase b to phosphorylase a in muscle extracts. *J. Biol. Chem.* 216, 121–132 (1955).

2 Deutscher J, Saier MH. Ser/Thr/Tyr protein phosphorylation in bacteria – for long time neglected, now well established. *J. Mol. Microbiol. Biotechnol.* 9, 125–131 (2005).

3 Hoch JA. Two-component and phosphorelay signal transduction. *Curr. Opin. Microbiol.* 3, 165–170 (2000).

4 Almengor AC, Kinkel TL, Day SJ, McIver KS. The catabolite control protein CcpA binds to P*mga* and influences expression of the virulence regulator Mga in the Group A *Streptococcus*. *J. Bacteriol.* 189, 8405–8416 (2007).

5 Liang W, Pascual-Montano A, Silva AJ, Benitez JA. The cyclic AMP receptor protein modulates quorum sensing, motility and multiple genes that affect intestinal colonization in *Vibrio cholerae*. *Microbiology* 153, 2964–2975 (2007).

6 Cozzone AJ. Role of protein phosphorylation on serine/threonine and tyrosine in the virulence of bacterial pathogens. *J. Mol. Microbiol. Biotechnol.* 9, 198–213 (2005).

7 Backert S, Selbach M. Tyrosine-phosphorylated bacterial effector proteins: the enemies within. *Trends Microbiol.* 13, 476–484 (2005).

8 Wehenkel A, Bellinzoni M, Grana M *et al.* Mycobacterial Ser/Thr protein kinases and phosphatases: physiological roles and therapeutic potential. *Biochim. Biophys. Acta* 1784, 193–202 (2008).

9 Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 241, 42–52 (1988).

10 Kennely PJ, Potts M. Life among the primitives: protein O-phosphatases in prokaryotes. *Front. Biosci.* 4, D372–D385 (1999).

11 Grangeasse C, Cozzone AJ, Deutscher J, Mijakovic I. Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology. *Trends Biochem. Sci.* 32, 86–94 (2007).

12 Mijakovic I, Musumeci L, Tautz L *et al.* *In vitro* characterization of the *Bacillus subtilis* protein tyrosine phosphatase YwqE. *J. Bacteriol.* 187, 3384–3390 (2005).

13 Morona JK, Morona R, Paton JC. Attachment of capsular polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for invasive disease. *Proc. Natl Acad. Sci. USA* 103, 8505–8510 (2006).

14 Ilan O, Bloch Y, Frankel G, Ullrich H, Geider K, Rosenshine I. Protein tyrosine kinases in bacterial pathogens are associated with virulence and production of exopolysaccharide. *EMBO J.* 18, 3241–3248 (1999).

15 Saskova L, Novakova L, Basler M, Branny P. Eukaryotic-type serine/threonine protein kinase StkP is a global regulator of gene expression in *Streptococcus pneumoniae*. *J. Bacteriol.* 189, 4168–4179 (2007).

16 Sharma K, Gupta M, Krupa A, Srinivasan N, Singh Y. EmbR, a regulatory protein with ATPase activity, is a substrate of multiple serine/threonine kinases and phosphatase in *Mycobacterium tuberculosis*. *FEBS J.* 273, 2711–2721 (2006).

17 Mougous JD, Gifford CA, Ramsdell TL, Mekalnos JJ. Threonine phosphorylation post-translationally regulates protein secretion in *Pseudomonas aerugionosa*. *Nat. Cell Biol.* 9, 791–803 (2007).

18 Arbibe L, Kim DW, Batsche E *et al.* An injected bacterial effector targets chromatin access for transcription factor NF-κB to alter transcription of host genes involved in immune responses. *Nat. Immunol.* 8, 47–56 (2007).

19 Hatekayama M. *Helicobacter pylori* CagA – a bacterial intruder conspiring gastric carcinogenesis. *Int. J. Cancer* 119, 1217–1223 (2006).

20 Walburger A, Koul A, Ferrari G *et al.* Protein kinase G from pathogenic mycobacteria promotes survival within macrophages. *Science* 304, 1800–1804 (2004).

21 Macek B, Mijakovic I, Olsen JV *et al.* The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. *Mol. Cell Proteomics* 6, 697–707 (2007).

•• **First global, gel-free and site-specific study of bacterial phosphorylation on serine/threonine/tyrosine residues, setting a new standard in bacterial phosphoproteomics.**

22 Macek B, Gnad F, Soufi B *et al.* Phosphoproteome analysis of *E. coli* reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Mol. Cell Proteomics* 7, 299–307 (2008).

• **With the emerging datasets on bacterial phosphorylation sites, this study offers an evolutionary perspective of bacterial protein phosphorylation.**

23 Voisin S, Watson DC, Tessier L *et al.* The cytoplasmic phosphoproteome of the Gram-negative bacterium *Campylobacter jejuni*: evidence for modification by unidentified protein kinases. *Proteomics* 7, 4338–4348 (2007).

24 Kang CM, Abbot DW, Park ST *et al.* The *Mycobacterium tuberculosis* serine/threonine kinases PknA and PknB: substrate identification and regulation of cell shape. *Genes Dev.* 19, 1692–1704 (2005).

25 Rajagopal L, Clancy A, Rubens CE. A eukaryotic type serine/threonine kinase and phosphatase in *Streptococcus agalactiae* reversibly phosphorylate an inorganic pyrophosphatase and affect growth, cell segregation, and virulence. *J. Biol. Chem.* 278, 14429–14441 (2003).

26 Zheng X, Papavinasasundaram KG, Av-Gay Y. Novel substrates of *Mycobacterium tuberculosis* PknH Ser/Thr kinase. *Biochem. Biophys. Res. Commun.* 355, 162–168 (2007).

27 Mijakovic I, Poncet S, Boel G *et al.* Transmembrane modulator-mediated regulation of UDP-glucose DH, the first substrate of a bacterial tyrosine kinase. *EMBO J.* 22, 4709–4718 (2003).

28 Mijakovic I, Petranovic D, Macek B *et al.* Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res.* 34, 1588–1596 (2006).

29 Petranovic D, Michelsen O, Zahradka K *et al. Bacillus subtilis* strain deficient for the protein-tyrosine kinase PtkA exhibits impaired DNA replication. *Mol. Microbiol.* 63, 1797–1805 (2007).

30 Jin H, Pancholi V. Identification and biochemical characterization of a eukaryotic-like serine/threonine kinase and its cognate phosphatase in *Streptococcus pyogenes*: their biological functions and substrate identification. *J. Mol. Biol.* 357, 1351–1372 (2006).

31 Kirstein J, Zühlke D, Gerth U, Turgay K, Hecker M. A tyrosine kinase and its activator control the activity of the CtsR heat shock repressor in *B. subtilis*. *EMBO J.* 24, 3435–3445 (2005).

32 Stephenson K, Hoch JA. Two-component and phosphorelay signal-transduction systems as therapeutic agents. *Curr. Opin. Pharmacol.* 2, 507–512 (2002).

33 Beier D, Gross R. Regulation of bacterial virulence by two-component systems. *Curr. Opin. Microbiol.* 9, 143–152 (2006).

• **Very insightful review offering an exhaustive overview of stimuli and effects operated by the two-component systems involved in bacterial virulence.**

34 Ohtani K, Kawsar HI, Okumura K, Hayashi H, Shimizu T. The VirR/VirS regulatory cascade affects transcription of plasmid-encoded putative virulence genes in *Clostridium perfrigens* strain 13. *FEMS Microbiol. Lett.* 222, 137–141 (2003).

35 Dong YH, Zhang XF, Soo HM, Greenberg EP, Zhang LH. The two-component response regulator PprB modulates quorum-sensing signal production and global gene expression in *Pseudomonas aerugionsa*. *Mol. Micorobiol.* 56, 1287–1301 (2005).

36 Mouslim C, Delgado M, Groisman EA. Activation of the RcsC/YojN/RcsB phosphorelay system attenuates *Salmonella* virulence. *Mol. Microbiol.* 54, 386–395 (2004).

37 Eguchi Y, Oshima T, Mori H *et al.* Transcriptional regulation of drug efflux genes by EvgAS, a two-component system in *Escherichia coli*. *Microbiology* 149, 2819–2828 (2003).

38 Rajagopal L, Vo A, Silvestroni A, Rubens CE. Regulation of cytotoxin expression by converging eukaryotic-type and two-component signaling mechanisms in *Sterptococcus agalactiae*. *Mol. Microbiol.* 62, 941–957 (2006).

39 Stephenson K, Hoch JA. Histidine kinase-mediated signal transduction systems of pathogenic microorganisms as targets for therapeutic intervention. *Curr. Drug Targets Infect. Disord.* 2, 235–246 (2002).

40 Farris M, Grant A, Richardson TB, O'Connor CD. BipA: a tyrosine-phosphorylated GTPase that mediates interactions between enteropathogenic *Escherichia coli* (EPEC) and epithelial cells. *Mol. Microbiol.* 28, 265–279 (1998).

41 Grant AJ, Farris M, Alefounder P, Williams PH, Woodward MJ, O'Connor CD. Co-ordination of pathogenicity island expression by the BipA GTPase in enteropathogenic *Escherichia coli* (EPEC). *Mol. Microbiol.* 48, 507–521 (2003).

42 Herro R, Poncet S, Cossart P *et al.* How seryl-phosphorylated HPr inhibits PrfA, a transcription activator of *Listeria monocytogenes* virulence genes. *J. Mol. Microbiol. Biotechnol.* 9, 224–234 (2005).

43 Mertins S, Joseph B, Goetz M *et al.* Interference of components of the phosphoenolpyruvate phosphotransferase system with the central virulence gene regulator PrfA of *Listeria monocytogenes*. *J. Bacteriol.* 189, 473–490 (2007).

44 Shelburne III SA, Keith D, Horstmann N *et al.* A direct link between carbohydrate utilization and virulence in the major human pathogen group A *Streptococcus*. *Proc. Natl Acad. Sci. USA* 105, 1698–1703 (2008).

45 Yang X, Kang CM, Brody MS, Price CW. Opposing pairs of serine protein kinases and phosphatases transmit signals of environmental stress to activate a bacterial transcription factor. *Genes Dev.* 10, 2265–2275 (1996).

46 Kazmierczak MJ, Wiedmann M, Boor KJ. Contributions of *Listeria monocytogenes* σB and PrfA to expression of virulence and stress response genes during extra- and intracellular growth. *Microbiology* 152, 1827–1838 (2006).

47 Kim H, Marquis H, Boor KJ. σB contributes to *Listeria monocytogenes* invasion by controlling expression of *inlA* and *inlB*. *Microbiology* 151, 3215–3222 (2005).

48 Pané-Farré J, Jonas B, Förstner K, Engelmann S, Hecker M. The σB regulon in *Staphylococcus aureus* and its regulation. *Int. J. Med. Microbiol.* 296, 237–258 (2006).

49 van Schaik W, van der Voort M, Molenaar D, Moezelaar R, de Vos WM, Abee T. Identification of the σB regulon of *Bacillus cereus* and conservation of σB-regulated genes in low-GC-content Gram-positive bacteria. *J. Bacteriol.* 189, 4384–4390 (2007).

50 Williams P. Role of the cell envelope in bacterial adaptation to growth *in vivo* in infections. *Biochimie* 70, 987–1011 (1988).

51 Collins RF, Beis K, Dong CJ *et al.* The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* 104, 2390–2395 (2007).

52 Bender MH, Cartee RT, Yother J. Positive correlation between tyrosine phosphorylation of CpsD and capsular polysaccharide production in *Streptococcus pneumoniae*. *J. Bacteriol.* 185, 6057–6066.

53 Obadia B, Lacour S, Doublet P *et al.* Influence of tyrosine-kinase Wzc activity on colanic acid production in *Escherichia coli* K12 cells. *J. Mol. Biol.* 367, 42–53 (2007).

54 Lacour S, Doublet P, Obadia B, Cozzone AJ, Grangeasse C. A novel role for protein-tyrosine kinase Etk from *Escherichia coli* K-12 related to polymyxin resistance. *Res. Microbiol.* 157, 637–641 (2006).

55 Grangeasse C, Obadia B, Mijakovic I, Deutscher J, Cozzone AJ, Doublet P. Autophosphorylation of the *Escherichia coli* protein kinase Wzc regulates tyrosine phosphorylation of Ugd, a UDP-glucose dehydrogenase. *J. Biol. Chem.* 278, 39323–39329 (2003).

56 Soulat D, Grangeasse C, Vaganay E, Cozzone AJ, Duclos B. UDP-acetyl-mannosamine dehydrogenase is an endogenous protein substrate of *Staphylococcus aureus* protein-tyrosine kinase activity. *J. Mol. Microbiol. Biotechnol.* 13, 45–54 (2007).

57 Novakova L, Saskova L, Pallova P *et al.* Characterization of a eukaryotic type serine/threonine protein kinase and protein phosphatase of *Streptococcus pneumoniae* and identification of kinase substrates. *FEBS J.* 272, 1243–1254 (2005).

58 Molle V, Brown AK, Besra GS, Cozzone AJ, Kremer L. The condensing activities of the *Mycobacterium tuberculosis* type II fatty acid synthase are differentially regulated by phosphorylation. *J. Biol. Chem.* 281, 30094–30103 (2006).

59 Zhao X, Lam JS. WaaP of *Pseudomonas aeruginosa* is a novel eukaryotic type protein-tyrosine kinase as well as a sugar kinase essential for the biosynthesis of core lipopolysaccharide. *J. Biol. Chem.* 277, 4722–4730 (2002).

60 Mattoo S, Lee YM, Dixon JE. Interaction of bacterial effector proteins with host proteins. *Curr. Opin. Immunol.* 19, 392–401 (2007).

61 Kulasekara HD, Miller SI. Threonine phosphorylation times bacterial secretion. *Nat. Cell Biol.* 9, 734–736 (2007).

62 Johnson DL, Mahony JB. *Chlamydophila pneumoniae* PknD exhibits dual amino acid specificity and phosphorylates Cpn0712, a putative type III secretion YscD homolog. *J. Bacteriol.* 189, 7549–7555 (2007).

• **Interesting from the standpoint of regulating pathogenicity at the level of effector secretion, but even more so from the evolutionary standpoint, as it is the first example of a bacterial protein kinase capable of phosphorylating serine, threonine and tyrosine residues.**

63 Kozak NA, Mattoo S, Foreman-Wykert AK, Whitelegge JP, Miller JF. Interactions between partner switcher orthologs BtrW and BtrV regulate type III secretion in *Bordetella*. *J. Bacteriol.* 187, 5665–5676 (2005).

64 Murli S, Watson RO, Galan JE. Role of tyrosine kinases and the tyrosine phosphatase SptP in the interaction of *Salmonella* with the host cells. *Cell. Microbiol.* 3, 795–810 (2001).

65 Bach H, Papavinasasundaram KG, Wong D, Hmama Z, Av-Gay Y. *Mycobacterium tuberculosis* virulence is mediated by PtpA dephosphorylation of human vacuolar protein sorting 33B. *Cell Host Microbe* 3, 316–322 (2008).

66 Underwood W, Zhang S, He SY. The *Pseudomonas syringae* type III effector tyrosine phosphatase HopAO1 suppresses innate immunity in *Arabidopsis thaliana*. *Plant J.* 52, 658–672 (2007).

67 Trasak C, Zenner G, Vogel A *et al. Yersinia* protein kinase YopO is activated by a novel G-actin binding process. *J. Biol. Chem.* 282, 2268–2277 (2007).

68 Alonso A, Bottini N, Bruckner S *et al.* Lck dephosphorylation at Tyr-394 and inhibition of T cell antigen receptor signaling by *Yersinia* phosphatase YopH. *J. Biol. Chem.* 279, 4922–4928 (2004).

69 Mukherjee S, Keitany G, Li Y *et al. Yersinia* YopJ acetylates and inhibits kinase activation by blocking phosphorylation. *Science* 312, 1211–1214 (2006).

• **Exciting study describing how a bacterial pathogen uses a transacetylase to acetylate and thereby 'mask' some specific phosphorylation sites in the target proteins, therefore blocking a human signaling cascade.**

70 Maresso AW, Deng Q, Pereckas MS *et al. Pseudomonas aeruginosa* ExoS ADP-ribosyltransferase inhibits ERM phosphorylation. *Cell. Microbiol.* 9, 97–105 (2007).

71 Selbach M, Moese S, Hurwitz R, Hauck CR, Meyer TF, Backert S. The *Helicobacter pylori* CagA protein induces cortactin dephosphorylation and actin rearrangement by c-Src inactivation. *EMBO J.* 22, 515–528 (2003).

72 Bommarius B, Maxwell D, Swimm A *et al.* Enteropathogenic *Escherichia coli* Tir is an SH2/3 ligand that recruits and activates tyrosine kinases required for pedestal formation. *Mol. Microbiol.* 63, 1748–1768 (2007).

73 Ido JW, Carlson AC, Kennedy EL. *Anaplasma phagocytophilum* AnkA is tyrosine-phosphorylated at EPIYA motifs and recruits SHP-1 during early infection. *Cell. Microbiol.* 9. 1284–1296 (2007).

74 Rahmsdorf HJ, Pai SH, Ponta H *et al.* Protein kinase induction in *Escherichia coli* by bacteriophage T7. *Proc. Natl Acad. Sci. USA* 71, 586–589 (1974).

75 Severinova E, Severinov K. Localization of the *Escherichia coli* RNA polymerase β´ subunit residue phosphorylated by bacteriophage T7 kinase Gp0.7. *J. Bacteriol.* 188, 3470–3476 (2006).

76 Gafken PR, Lampe PD. Methodologies for characterizing phosphoproteins by mass spectrometry. *Cell Commun. Adhes.* 13, 249–262 (2006).

• **Concise overview of various separation and identification techniques currently used in the detection of phosphoproteins by mass spectrometry. Excellent introduction for newcomers to the field.**

77 Collins MO, Yu L, Choudhary JS. Analysis of protein phosphorylation on a proteome-scale. *Proteomics* 7, 2751–2768 (2007).

78 Sweet SM, Cooper HJ. Electron capture dissociation in the analysis of protein phosphorylation. *Expert Rev. Proteomics* 4, 149–159 (2007).

79 Aebersold A, Mann M. Mass spectrometry-based proteomics. *Nature* 422, 198–207 (2003).

80 Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 1, 252–262 (2005).

• **Excellent overview of different methods and state of the art in quantitative proteomics. Includes useful tips on when/how to use quantitative proteomics and how to choose the appropriate technique.**

81 De Leenher AP, Thienpont LM. Applications of isotope dilution-mass spectrometry in clinical chemistry, pharmcokinetics, and toxicology. *Mass Spectrom. Rev.* 11, 249–307 (1992).

82 Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* 17, 994–999 (1999).

83 Ong SE, Mann M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* 1, 2650–2660 (2006).

•• **Seminal paper describing the nonradioactive isotope labeling method that underlies most of the currently employed quantitative proteomics methods.**

84 Mann M. Functional and quantitative proteomics using SILAC. *Nat. Rev. Mol. Cell Biol.* 7, 952–958 (2006).

85 Krüger M, Kratchmarova I, Blagoev B, Tseng YH, Kahn CR, Mann M. Dissection of the insulin signaling pathway via quantitative phosphoproteomics. *Proc. Natl Acad. Sci. USA* 105, 2451–2456 (2008).

• **Recent study that convincingly demonstrates how quantitative proteomics can be used to dissect a very complex signaling pathway.**

86 Navaza AP, Encinar JR, Carrascal M, Abian J, Sanz-Medel A. Absolute and site-specific quantification of protein phosphorylation using integrated elemental and molecular mass spectrometry: its potential to assess phosphopeptide enrichment procedures. *Anal. Chem.* 80(5), 1777–1787 (2008).

87 Linding R, Jensen LJ, Ostheimer GJ *et al.* Systematic discovery of *in vivo* phosphorylation networks. *Cell* 129, 1415–1426 (2007).

88 Shlomi T, Eisenberg Y, Sharan R, Ruppin E. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Mol. Syst. Biol.* 3, 101 (2007).

89 Schilling O, Frick O, Herzberg C *et al.* Transcriptional and metabolic responses of *Bacillus subtilis* to the availability of organic acids: transcription regulation is important but not sufficient to account for metabolic adaptation. *Appl. Environ. Microbiol.* 73, 499–507 (2007).

Affiliations

- Carsten Jers, MSc
Technical University of Denmark, Biosys, Center for Microbial Biotechnology, DK-2800 Lyngby, Denmark
Tel.: +45 4525 2495
Fax: +45 4593 2809
cj@bio.dtu.dk
- Boumediene Soufi, MSc
Technical University of Denmark, Biosys, Center for Microbial Biotechnology, DK-2800 Lyngby, Denmark
Tel.: +45 4525 2495
Fax: +45 4593 2809
bs@bio.dtu.dk
- Christophe Grangeasse, PhD
Institute of Biology and Chemistry of Proteins, CNRS, University of Lyon, 7 passage du Vercors, 69367 Lyon, France
Tel.: +33 472 722 688
Fax: +33 472 722 601
c.grangeasse@ibcp.fr
- Josef Deutscher, PhD
Laboratory of Microbiology and Molecular Genetics, CNRS, AgroParisTech, INRA, 78850 Thiverval-Grignon, France
Tel.: +33 130 815 447
Fax: +33 130 815 457
josef.deutscher@grignon.inra.fr
- Ivan Mijakovic, PhD
Technical University of Denmark, Biosys, Center for Microbial Biotechnology, DK-2800 Lyngby, Denmark
Tel.: +45 4525 2495
Fax: +45 4593 2809
im@bio.dtu.dk

DOI 10.1002/pmic.200800285 *Proteomics* 2009, *9,* 116–125

RESEARCH ARTICLE

NetPhosBac – A predictor for Ser/Thr phosphorylation sites in bacterial proteins*

*Martin Lee Miller[1]**, ***, Boumediene Soufi[2]**, Carsten Jers[2], Nikolaj Blom[1], Boris Macek[3] and Ivan Mijakovic[2]*

[1] Center for Biological Sequence Analysis, Technical University of Denmark, Lyngby, Denmark
[2] Center for Microbial Biotechnology, Technical University of Denmark, Lyngby, Denmark
[3] Department of Proteomics and Signal Transduction, Max-Planck-Institute for Biochemistry, Martinsried, Germany

Received: April 1, 2008
Revised: July 11, 2008
Accepted: July 15, 2008

There is ample evidence for the involvement of protein phosphorylation on serine/threonine/tyrosine in bacterial signaling and regulation, but very few exact phosphorylation sites have been experimentally determined. Recently, gel-free high accuracy MS studies reported over 150 phosphorylation sites in two bacterial model organisms *Bacillus subtilis* and *Escherichia coli*. Interestingly, the analysis of these phosphorylation sites revealed that most of them are not characteristic for eukaryotic-type protein kinases, which explains the poor performance of eukaryotic data-trained phosphorylation predictors on bacterial systems. We used these large bacterial datasets and neural network algorithms to create the first bacteria-specific protein phosphorylation predictor: NetPhosBac. With respect to predicting bacterial phosphorylation sites, NetPhosBac significantly outperformed all benchmark predictors. Moreover, NetPhosBac predictions of phosphorylation sites in *E. coli* proteins were experimentally verified on protein and site-specific levels. In conclusion, NetPhosBac clearly illustrates the advantage of taxa-specific predictors and we hope it will provide a useful asset to the microbiological community.

Keywords:
Artificial neural network / Bacteria / Kinase / Prediction / Protein phosphorylation

1 Introduction

Protein phosphorylation is a widespread and abundant PTM, which is implicated in regulation and signal transduction [1]. It is estimated that 30–40% of all proteins in a eukaryotic cell are phosphorylated [2] making phosphorylation-dependent signaling networks complex, and rendering the assignment of phosphorylation sites to their proper kinases and phosphatases a challenging task [3]. Traditionally, bacterial signal transduction and regulation was perceived as mainly dependent on histidine/aspartate phosphorylation, which figures prominently in two-components systems and regulatory phenomena related to sugar transport [4–6]. However, evidence concerning the presence of serine/threonine/tyrosine phosphorylation in bacteria accumulated over the years [7–9], and today it is clear that this type of protein phosphorylation plays an equally prominent role in bacterial physiology [10, 11]. Bacteria do possess a number of eukaryotic-like protein kinases [8], but there is also clear evidence for idiosyncratic bacterial kinases that have evolved independently [11–13].

Correspondence: Dr. Ivan Mijakovic, Center for Microbial Biotechnology, Technical University of Denmark, DK-2800 Lyngby, Denmark
E-mail: im@biocentrum.dtu.dk
Fax: +45-45932809

Abbreviations: ANN, artificial neural network; **AROC,** area under receiver operating characteristics curve; **ROC,** receiver operating characteristics

* The software is available at: http://www.cbs.dtu.dk/services/NetPhosBac
** Both these authors contributed equally to this work
*** Additional corresponding author: Martin Lee Miller, E-mail: miller@cbs.dtu.dk

This explains the fact that currently available protein phosphorylation site predictors, largely based on eukaryotic datasets [14–16], are not ideally suited for prediction of bacterial phosphorylation sites [17]. On the other hand, construction of bacterial-specific predictors has been hindered by very limited datasets of bacterial phosphorylation sites [18]. In that respect, the new generation of phosphoproteomics studies based on the gel-free/high accuracy MS approach signaled a major breakthrough. Recently performed studies on *Bacillus subtilis* [19] and *Escherichia coli* [20] provided a valuable basis for bacterial phosphorylation predictors, yielding a total of over 150 determined phosphorylation sites. These bacterial datasets were combined with a literature mined dataset and used to train an artificial neural network (ANN)-based predictor, NetPhosBac. The performance of NetPhosBac in predicting bacterial phosphorylation sites was significantly better compared to other available predictors. Moreover, the high prediction accuracy of the predictor was demonstrated by experimental confirmation in *E. coli*.

2 Materials and methods

2.1 Dataset and handling

We obtained serine and threonine phosphorylation sites from the following sources: 14 sites from the Phosphorylation Site Database [18], 71 sites from *B. subtilis* [19] and 102 sites from *E. coli* [20] (See Table S1 of Supporting Information). Serine and threonine-containing peptides detected during the MS runs without any detected phospho-groups were used as negatives. Homology reduction was first performed on full protein length and subsequently on 13-mer peptide level (positive or negative site centrally placed) using CD-HIT [21] with default values and 90% sequence identity threshold in both cases. After homology reduction, we randomly downsampled the negative set to include only five negative examples *per* positive example, since machine learning methods often work poorly if applied to unbalanced datasets that differ in orders of magnitude.

2.2 Artificial neural network training

The ANN used in this work was a standard three-layer feed forward type that has been described previously [22]. *N*-fold crossvalidation was applied as described previously [23]. In this extended *N*-fold crossvalidation approach, an external validation set that has not been used to optimize any parameter (including when to stop training) is used to report the final performance of the ANN. Briefly, we used fourfold crossvalidation by training the ANN on two subsets (training), determining the optimal parameters on the third subset (test) and obtaining an unbiased performance estimation from the fourth subset (validation). The subsets were then shuffled in a round robin fashion. The following parameters were varied to optimize Matthews correlation coefficient on each test set: the window size (5, 7, 9, 11, and 13) and hidden neurons (0, 4, 8, and 12). The learning rate was 0.05. The best parameters from each crossvalidation were kept in an ensemble of 12 synapse files, which constituted the final ANN. The optimal window size and number of hidden neurons chosen were 11 and 8, respectively.

2.3 Benchmarking

The collected phosphorylation site dataset (see Date set and handling) was run through DisPhos [24] (ticking the bacterial option on), NetPhos [22], and NetPhosYeast [23] prediction resources and a receiver operating characteristics (ROC) curve was made for each of the methods. Furthermore, for NetPhosBac a ROC curve was also calculated based on the independent crossvalidation sets (see above). The performances of DisPhos (Bacteria), NetPhos, NetPhosYeast were compared to NetPhosBac as judged by resampling the scores of positive and negative examples to construct a bootstrapped area under ROC curve (AROC) distribution.

2.4 Evaluation of NetPhosBac on phosphoproteins

Non site-specific phosphoproteins identified in *B. subtilis* [25, 26] and *Corynebacterium glutamicum* [9] and the background proteomes pruned for the phosphoproteins were scored with NetPhosBac. In *B. subtilis* 228 out of 1324 sites in the phosphoproteins scored higher than 0.5 compared to 26 285 out of 132 899 in the background proteome. In *C. glutamicum* 300 out of 1650 sites in the phosphoproteins scored higher than 0.5 compared to 23 225 out of 118 674 in the background proteome. To determine the possible enrichment of high scoring sites in the phosphoproteins compared to the background proteome, hypergeometric sampling without replacement was performed.

2.5 Sequence logos

The *E. coli* and *B. subtilis* serine and threonine phosphopeptides were cut into 13-mers with the phosphorylation site centrally placed. Only unique phosphopeptides were considered. Using the negative phosphopeptides as background frequencies, logo plots were made using the EnoLogo program [27]. The EnoLogo program is based on sequence logos, which is a graphical method for displaying patterns in aligned sequences [28]. We collected a eukaryotic phosphorylation site dataset from phospho.ELM database [29], which consists of phosphorylation sites in mainly mammalian species, and phosphorylation sites gathered in a study of the yeast phosphoproteome [23]. After pruning tyrosine phosphorylation sites, cutting out 13-mers and reducing homologous peptides that were more than 90% identical, we obtained 11 227 serine and threonine eukaryotic phospho-

peptides. These were compared to the dataset of prokaryotic phosphorylation sites using two sample logo plot with default parameter settings [30].

2.6 Selecting proteins for phosphorylation assay

The *E. coli* proteome was pruned from proteins that were previously identified as phosphorylated by MS [20]. Furthermore, SignalP [31] for Gram-negative bacteria with default settings (truncating on 70 amino acids as well as using both ANN and HMM) and TMHMM [32] with default parameters were used to prune secreted and transmembrane proteins, respectively. This pruning resulted in a reduction of proteins from 4302 to 2442 candidates, which were subsequently run through the NetPhosBac predictor. The four highest-scoring purifiable proteins (Nfi, CitF, NudL, and GlcC) and lowest-scoring purifiable proteins (YhaL, HypC, YjiS, and YjiX) on single serine or threonine residues were selected for *in vivo* phosphorylation assay.

2.7 Strains and growth conditions

E. coli NM522 was used for plasmid propagation in the cloning experiments. The chaperone overproducing strain *E. coli* M15 carrying pREP4-GroESL [33] was used for overproduction of protein for the *in vivo* phosphorylation assay. Cells were grown at 37°C with vigorous shaking in Luria–Bertani medium supplemented with 100 μg/mL ampicillin. For M15, the medium was further supplemented with 25 μg/mL kanamycin.

2.8 DNA manipulations

The genes *fruB*, *nfi*, *citF*, *nudL*, *glcC*, *yaiL*, *hypC*, *yjiS*, and *yjiX* were PCR-amplified using *E. coli* K-12 MG1655 genomic DNA as template and specific primers with appropriate restriction sites added (see Table 1). PCR products were inserted between the *Bam*HI and *Hind*III restriction sites in the vector pQE30 (Qiagen) and the constructed vectors were propagated in *E. coli* NM522. The inserts were sequenced to verify that no mutations were introduced by PCR.

2.9 *In vivo* phosphorylation assay

Strains derived from *E. coli* M15 pREP4-GroESL, and carrying a specific gene inserted in pQE-30, were grown in 1 L of LB medium. Synthesis of 6 × His-tagged proteins was induced with 1 mM IPTG at OD 0.45 and cells were harvested 3 h after induction. Cell pellets were incubated in 5 mL solution A (50 mM Tris-Cl, pH 7.5, 100 mM NaCl, 10% glycerol) with 1 mg/mL lysozyme, and 5 μg/mL DNase for 15 min at 37°C. Protoplasts were disrupted by sonication and the cell debris was precipitated by centrifugation at 25 000 × *g* for 15 min. 6 × His-tagged proteins were purified from crude extracts with Ni-NTA columns as described by the manufacturer (Qiagen), and purified proteins were desalted on PD-10 columns (GE Healthcare).

Proteins were separated by electrophoresis on 12% SDS-polyacrylamide gels in duplicate. One gel was stained with CBB G-250 whereas the second was Western blotted with antiphosphoS/T antibody. Briefly, SDS-polyacrylamide gel was soaked in 100 mL of Tris-Glycine buffer (25 mM Tris base, 192 mM glycine, 10% ethanol) for 5 min. Proteins were electroblotted on a PVDF membrane (BioRad). Membrane was blocked overnight at 4C in PBS buffer containing 3% BSA fraction V and 0.05% Tween 20. Blocked membrane was incubated with the primary antibody, monoclonal PSR-45 (Sigma), diluted 1:1000 for 60 min at room temperature, then washed 3 × for 10 min with 0.05% Tween 20 in PBS. This was followed by incu-

Table 1. PCR primers used for amplification with underlined restriction sites

nfi forward	CGCGGATCCATGGATCTCGCGTCATTACGCGCTC	*Bam*HI
nfi reverse	CATTAAGCTTTTAGGGCTGATTTGCTGTATAGCGC	*Hind*III
citF forward	CGCGGATCCATGACGCAGAAAATTGAACAATCTC	*Bam*HI
citF reverse	CATTAAGCTTTTATTCCTTCACCTGATGCACAACATC	*Hind*III
nudL forward	CGCGGATCCGTGGAATACCGTAGCCTGACGCTTG	*Bam*HI
nudL reverse	CATTAAGCTTTCAGGGTTTCACACCAATTTGCAGC	*Hind*III
glcC forward	CGCGGATCCATGAAAGATGAACGTCGCCCTATTTG	*Bam*HI
glcC reverse	CATTAAGCTTCTAACTCAGGTTCATCTCCAGCGGC	*Hind*III
yhaL forward	CGCGGATCCATGAGTAAAAAATTGGCCAAAAAGC	*Bam*HI
yhaL reverse	CATTAAGCTTTTACGCCGTAGCTTCATCCTCGGC	*Hind*III
hypC forward	CGCGGATCCATGTGCATAGGCGTTCCCGGCCAG	*Bam*HI
hypC reverse	CATTAAGCTTTTATTTTTCCTCGCCATACAAC	*Hind*III
yjiS forward	CGCGGATCCATGGAATTTCACGAAAACAGAGC	*Bam*HI
yjiS reverse	CATTAAGCTTTCACTCCACATCCTCCCTGCGTAAG	*Hind*III
yjiX forward	CGCGGATCCATGTTTGGTAACTTAGGACAGGC	*Bam*HI
yjiX reverse	CATTAAGCTTTTAACAACAGCGCATACCGCC	*Hind*III

bation with the secondary antibody, antibiotin peroxidase conjugate (Sigma), diluted 1:4000 for 45 min at room temperature ensued by 3 × 10 min washes with 0.05% Tween 20 in PBS. Bound secondary antibody was detected with the AEC staining kit (Sigma).

2.10 MS analysis of phosphoproteins

Lyophilized purified proteins (amounts 100–400 μg) were dissolved in denaturing solution (6 M urea, 2 M thiourea in 20 mM Tris, pH 7.0), to the concentration of 1 μg/μL. Proteins were reduced in 1 mM dithiotreithol for 45 min at room temperature and carbamidomethylated with 5 mM iodoacetamide for 45 min in the dark. Protein digestion was performed with endoprotease Lys-C (Waco) for 3 h at room temperature, followed by dilution with four volumes of 20 mM ammonium bicarbonate and further overnight digestion with sequence grade modified trypsin (Promega). The enzyme/protein ratio was in both cases 1/50. Trypsin activity was quenched by addition of TFA to the final concentration of 0.01%. Resulting digests were subjected to phosphopeptide enrichment using TiO_2 beads (GL Sciences). Prior to this step, ACN was added to all samples to the final concentration of 30% and the TiO_2 beads were preincubated with a solution containing 20 mg/mL dihydroxybenzoic acid in 80% ACN. TiO_2 beads slurry containing 5 mg of beads was added to each sample and incubated for 1 h with end-overend rotation. Beads were washed twice in a solution containing 60% ACN and 0.1% TFA and eluted with 150 μL of 15%-ammonia solution in 60% ACN (pH 10.5). Eluates were prepared for LC-MS measurements by reducing their volumes to 5 μL and adding equal volume of a solvent containing 2% ACN and 1% TFA.

MS was performed on a LTQ-Orbitrap XL mass spectrometer (Thermo), coupled to an Agilent 1100 nano-HPLC. Enriched phosphopeptides were separated on a home-made 75 μm id nano-HPLC column emitter, packed with 3 μm C18 beads (Dr. Maisch), using a 100 min. linear gradient of solvent "B" (80% ACN, 0.5% acetic acid). Mass spectrometer was operated in the positive ion mode, under "Top-10" acquisition regime, consisting of a MS survey scan in the orbitrap analyzer (at resolution of 60 000), followed by MS/MS of ten most intense peptide ions in the linear IT analyzer (LTQ). Lock mass to the background syloxane ions was used to improve mass accuracy, and resulting spectra were processed using MaxQuant software (v.1.0.11.5) and searched using MASCOT search engine (Matrix Science), against concatenated human and *E. coli* decoy protein database containing 147 268 entries. MASCOT search criteria were: required trypsin specificity, allowed two missed cleavages, precursor ion mass tolerance 7 ppm, fragment ion mass tolerance 0.5 Da, variable modifications oxidation (M), N-acetylation (Protein), and Phosphorylation (STY); carbamidomethylation (C) was considered as fixed modification.

3 Results and discussion

3.1 Bacterial phosphorylation sites

In bacteria, two component histidine kinases that phosphorylate cognate response regulators on aspartate were traditionally believed to comprise the most of phosphorylation-dependent signaling [5]. Sequenced bacterial genomes encode anywhere between 20 and 42-component systems, depending on the lifestyle of the given bacterium. However, it recently became obvious that bacterial proteins phosphorylated on serine, threonine, and tyrosine residues largely outnumber the two-component systems, with over 100 phosphorylation sites detected *per* bacterial cell [19, 20]. To obtain a comprehensive bacterial phosphorylation dataset, we collected sites described in literature and sites mapped by global-scale phosphoproteomics studies. Since the majority of the phosphorylation incidences in prokaryotes are on serine and threonine, we focused on developing a predictor for these residues. In total, we obtained 14 sites from the Phosphorylation Site Database [18], which is the main repository for prokaryotic phosphorylation sites, 78 sites from a phosphoproteomic study on the Gram-positive model organism *B. subtilis* [19] and 109 sites from a phosphoproteomic study on the Gram-negative model organism *E. coli* [20].

Kinase–substrate recognition is partly mediated through *linear motifs*, which are small unstructured sequence patches flanking PTMs such as phosphorylation sites [34]. To investigate the similarities between the sequence motifs in the two major components of the dataset, we aligned the *B. subtilis* and the *E. coli* sequences in a window of 13 residues around the modified sites. Subsequently, logo plots were used to visualize conserved residues in the vicinity of the modified sites (Figs. 1A and B). In logo plots, the height of the column represents the degree of conservation at that position, while the height of the individual letters is proportional to the relative frequency of the amino acid residue [28]. Only unique phosphopeptides were considered in the logo plots.

The phosphorylation sequence motifs of the two species have several characteristics in common. Apart from the centrally aligned phosphorylated serine and threonine residues, the positions −1 and +1 have the highest information content and presumably play an important role in the kinase–substrate recognition. Furthermore, hydrophobic residues both N and C-terminally to the phosphorylation site are conserved in the two species. Since the logo profiles did not differ significantly between the two species it seemed reasonable to pool the data and develop a single predictor for bacterial phosphorylation sites.

Recently, it has been shown that bacterial phosphoproteins and their phosphorylated residues are evolutionary more conserved in orthologous proteins of archaeal and eukaryotic species than their nonphosphorylated counterparts [20]. To investigate this on a sequence motif level, we used a two sample logo plot [30] to visualize the differences between the prokaryotic and eukaryotic phosphorylation

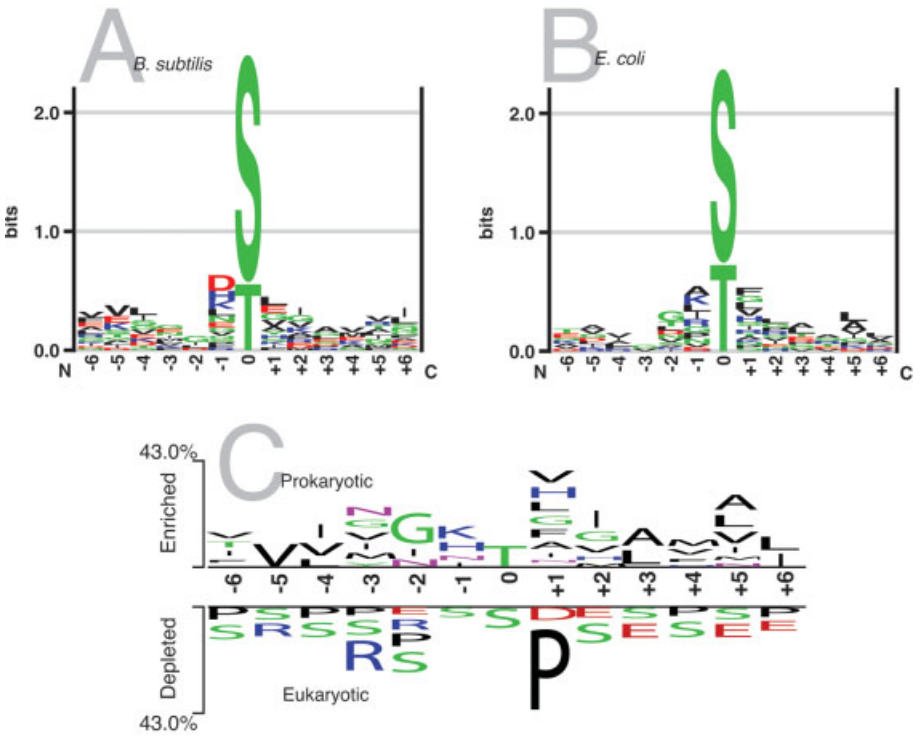


Figure 1. Logo plots of the (A) *B. subtilis* and (B) *E. coli* phosphoproteomes and (C) two sample logo plot of prokaryotic *versus* eukaryotic phosphoproteomes. Information content is shown in bits. The degree of conservation ranges from zero (no conservation) to approximately 4.3 bits (full conservation when all amino acids are equally probable). In the two-sample logo plot the "enriched" and "depleted" refers to the statistically significant differences in the amino acids between the two datasets. Position "0" is the position of the phosphorylated residue. Note that in both bacterial species phosphorylated peptides are characterized by hydrophobic amino acids (black), which is also evident when comparing the combined bacterial dataset with eukaryotic phosphorylation sites. The amino acids are colored according to their properties: Acidic [DE] red, basic [HKR] blue, hydrophobic [ACFILMPVW] black and neutral [GNQSTY] green. Only unique phosphopeptides were considered.

motifs (Fig. 1C). The eukaryotic serine and threonine phosphoproteome was collected by merging the yeast [23] and mammalian [29] phosphoproteomes. In the two-sample logo plot only significantly ($p<0.05$) enriched and depleted amino acids between the two input samples are shown.

Confirming the observation that hydrophobic N and C-terminal residues are a hallmark of bacterial phosphorylation sites, a similar enrichment of such amino acids was found when comparing prokaryotic to eukaryotic phosphosites. Unlike eukaryotic kinases, most bacterial kinase motifs have not yet been characterized. Some of the best characterized serine/threonine kinases in eukaryotes have very distinct motifs ranging from C-terminal basic amino acids (PKA, PKC, PKG, and CaMKII), proline at position +1 (CDKs) to acidic residues (CK1, CK2, and TLK) [14, 35, 36]. As evident from the two-sample logo plot, such motifs are depleted in bacteria, which is not surprising since there are rather few eukaryotic-like kinases in prokaryotes; for example, there are only around half a dozen such kinases in both *B. subtilis* and *E. coli* [37]. However, the visualized sequence motifs for *B. subtilis* and *E. coli* (Figs. 1A and B) represent a mix of substrates for different kinases.

3.2 Development of NetPhosBac

An essential step in training any method for predicting linear sequence motifs is data handling. Since similar phosphorylation sites in similar proteins cannot be considered as independent observations and because machine-learning methods developed on highly homologous data are susceptible to overfitting, we performed a conservative homology reduction. Homology reduction was made both on the level of the phosphorylation sites and on the full length of the proteins. After homology reduction the final dataset consisted of 103 serine and 37 threonine phosphorylation sites.

Machine-learning methods require both a positive dataset, such as the dataset described above, and a negative dataset. Assigning the negative data often imposes a considerable challenge, for example the physical and chemical properties of the negative proteins may not reflect those of the positive proteins. Consequently, the predictor will learn to discriminate between such differences rather than those due to the phosphorylation event. However, the main problem is that negative information is most often not available, since it is traditionally uncommon to report a protein as nonphos-

phorylated. For this reason other phospho-predictors use random S/T/Y residues as negatives [14–16, 38–41]. Therefore, there is a risk having false negatives in the training data, since the randomly selected residues might indeed be phosphorylated. In this work, we have the opportunity to define the negatives as those identified as nonphosphorylated by MS, since we compile the negatives from the nonphosphorylated peptides detected in the two phosphoproteomics studies. Using this approach, the risk of having false negatives in our negative dataset is minimized and we obtain the same relative amino acid frequencies in the positive and negative datasets. However, one cannot rule out the possibility that a "negative" protein could be phosphorylated under other growth conditions.

Machine-learning methods are able to account for interdependencies between residues at different positions and can thus capture the complex mechanisms that guide interactions such as kinase–substrate recognition [42]. Although many machine-learning algorithms are available, we chose to use ANNs as we have successfully used them to predict PTMs in previous studies [14, 22]. For ANN training we used an extension of the traditional *N*-fold crossvalidation procedure as described previously [23]. Here, an external evaluation set that has not been used to optimize any parameters (such as determining when to stop training) was used to evaluate the performance of the ANN.

The final predictor, named NetPhosBac, was constituted of an ensemble of 12 configuration (synapse) files, optimized to a specific set of ANN parameters, namely window size and number of hidden neurons. The average window size and number of hidden neurons that captures the sequence characteristics best is 12 and 9, respectively.

3.3 Performance and benchmark of NetPhosBac

Existing predictors for protein phosphorylation are primarily based on mammalian data, with NetPhos being the first machine learning-based method [22]. Later, predictors such as DisPhos have been trained to identify phosphorylation sites in different organisms and functional categories, although organism-specific phosphorylation sites were not used [24]. Recently, a species-specific predictor trained on yeast phosphorylation sites was released (NetPhosYeast) [23]. As more phosphorylation sites have been annotated to be substrate of a particular kinase and collected in public databases, numerous kinase-specific predictors have been developed [14–16, 38–41].

In order to compare NetPhosBac to existing predictors we used AROC as a performance measure. The AROC is an appropriate performance measure for binary classification, such as phosphorylation prediction, since it is not necessary to choose an arbitrary threshold for defining if a score signifies a positive or negative prediction.

Since no predictor has so far been trained on bacterial phosphorylation sites, we wanted to investigate how predictors trained on eukaryotic data perform on the bacterial phosphorylation data. We chose to use NetPhos, NetPhos Yeast, and DisPhos, since these are kinase-unspecific predictors and are consequently not biased toward specific kinase motifs. NetPhos and NetPhosYeast are based on ANNs, while DisPhos uses a linear model, position specific scoring matrices, combined with disorder prediction to discriminate between potentially phosphorylated and nonphosphorylated sites. The DisPhos prediction output can furthermore be adjusted to different classes of proteins, among these bacterial proteins, hence this option was chosen in the benchmark.

NetPhosBac obtains an AROC curve of 0.74 and significantly ($p<0.001$) outperforms NetPhos, NetPhosYeast, and DisPhos, which all have random performances (AROC curve ~0.5, see Fig. 2). However, this is expected since prokaryotic phosphorylation sites do not resemble eukaryotic ones (see Fig. 1C). The significant dissimilarities in the substrate specificities of kinases between different species, as illustrated here, clearly underline the need to use newly accumulated data to develop taxon-specific predictors.

Moreover, we were able to test the performance of NetPhosBac on an external dataset, which had not been used in the crossvalidation procedure. At the time we obtained the datasets from the two phosphoproteomics studies in *B. subtilis* and *E. coli* there were ten phosphorylation sites that could not unambiguously be assigned to a particular residue and consequently these were omitted in the training procedure. However, after manual inspection and reanalysis of these sites, at the time of publication they were reported in the two studies, giving the opportunity to use these as an external

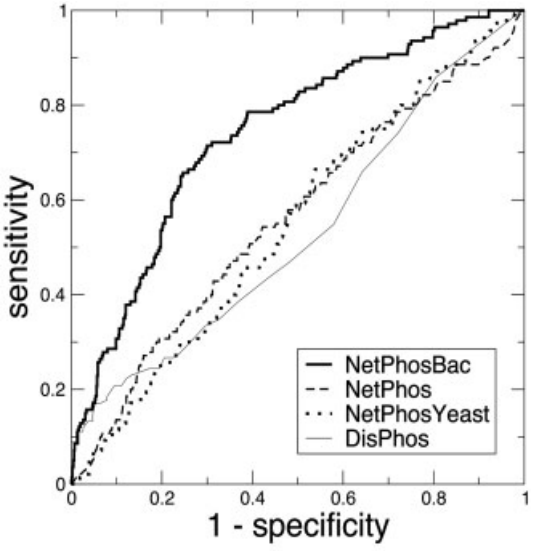


Figure 2. Performance of NetPhosBac and DisPhos. ROC curves are made by altering thresholds for classifying a score as a positive or negative and calculating the "true positive rate" (sensitivity) and the "false positive rate" (1-specificity) for each threshold. The AROC curve of 0.5 corresponds to a random performance. The obtained AROC values on the collected bacterial phosphorylation dataset were: NetPhosBac = 0.74; DisPhos = 0.53; NetPhos = 0.56; NetPhosYeast = 0.55.

validation set. In accordance with the performance obtained in the crossvalidation and benchmark procedures, NetPhosBac correctly predict nine out of the ten phosphorylation sites (see Table 2), while NetPhos, NetPhosYeast, and DisPhos predicts three, three, and two, respectively (results not shown).

Finally, we evaluated the performance of NetPhosBac on phosphoproteins identified in *B. subtilis* [25, 26] and *C. glutamicum* [9]. Since these studies did not investigate which particular residues were phosphorylated, we investigated if the phosphoproteins were enriched for high scoring (NetPhosBac score >0.5) serine and threonine residues compared to the background proteome in the two organisms (see methods for details). Indeed we find that this was the case both in *B. subtilis* ($p<0.002$) and in *C. glutamicum* ($p<0.01$).

3.4 Experimental verification of NetPhosBac predicted proteins

One very important issue with any computationally based prediction platform is to experimentally validate the performance and accuracy of the predictor. Primary literature sources for most available phosphorylation prediction tools did not report experimental verifications and we considered it appropriate to include such validation when reporting a new predictor. We decided to perform an experimental verification of NetPhosBac, using a different technique from the one used to create the training dataset (MS). We chose to use antiphosphoserine/threonine antibodies, focusing specifically on the phosphorylation of residues that NetPhosBac predicts. The specificity of the antibodies toward phosphorylated serine and threonine residues eliminates the possibility of detection of phospho-histidine/aspartate/tyrosine residues that are also present in bacterial systems.

We first subjected the entire proteome complement of *E. coli* to NetPhosBac to determine which proteins scored highest for phosphorylation. Prior to NetPhosBac predictions, those proteins that were previously identified by MS as positively phosphorylated were removed from the prediction dataset. Furthermore, SignalP [31] for Gram-negative bacteria and TMHMM [32] were used to prune secreted and transmembrane proteins (see Section 2), respectively, which would not be purifiable under our experimental conditions.

Four high-scoring candidates from *E. coli* (GlcC, NudL, Nfi, and CitF) and four low scoring candidates (YhaL, HypC, YjiS, and YjiX) were selected for antiphospho-S/T Western blot. These candidates have not previously been shown to be phosphorylated on S/T residues. In our assay, all high scoring candidates tested positive for phosphorylation, whereas among the predicted negatives three out of four tested negative for S/T phosphorylation (Fig. 3).

One might ask why these phosphoproteins were not previously identified by MS [20], and there are several potential explanations. In order to facilitate the purification of these proteins we used a chaperone-overproducing *E. coli* strain (M15 pREP4-GroESL), slightly different from the strain used in the MS study. Moreover, in our experiment the proteins were overproduced, and it is possible that under standard conditions their levels were insufficient for detection by MS. We have to acknowledge that no matter how powerful the present day analytical MS methods are, their detection power is not absolute and it is quite possible that the bacterial cells contain more phosphoproteins than we were able to detect [20]. The fact that one negative control (HypC) appeared to be phosphorylated in our assay might indicate that the training set did not account for substrates of all bacterial kinases, which is not surprising since it was derived from a single experimental condition, whereas protein phosphorylation is a dynamic event [25] (*i.e.*, largely dependent on growth conditions). The phosphorylation events that we identified here for the first time are interesting *per se*, since they might exert some regulatory function of the target proteins.

Table 2. Prediction of peptides with phosphorylation sites that could not unambiguously be assigned to a particular residue

Gene name	Organism	Accession number	Ambiguous peptide	Residue (position)	Prediction outcome
pgm	*E. coli*	1786904	GIVITPSHNPPED	S (146)	Correct
ptsI	*E. coli*	1788756	TDAGGRTSHTSIM	T (187)	Correct
sera	*E. coli*	1789279	IGLRSRTHLTEDV	T (63)	Correct
deoB	*E. coli*	1790843	SGKDTPSGHWEIA	S (100)	Correct
yjjJ	*E. coli*	1790845	GEIVGSSAGGEQP	S (201)	Incorrect
gatY	*E. coli*	87082041	LVLHGASGLSTKD	S (211)	Correct
ybbT	*B. subtilis*	Bsu0177	GVMISASHNPVQD	S (100)	Correct
rsbV	*B. subtilis*	Bsu0471	DVSYMDSTGLGVF	S (56)	Correct
yojH	*B. subtilis*	Bsu1945	LIGDIDTVRAKLI	T (186)	Correct
ypoC	*B. subtilis*	Bsu2232	QELREKSYPAKPI	S (117)	Correct

At the time of obtaining the datasets from the two phosphoproteomics studies of *B. subtilis* [19] and *E. coli* [20] there were ten peptides that were found to be phosphorylated but the phosphorylation site could not unambiguously be assigned to a particular residue. However, after manual inspection and reanalysis of these sites the specific positions were identified and they were reported in the two studies, giving the opportunity to use NetPhosBac to distinguish which of the ambiguous sites scored highest for phosphorylation.

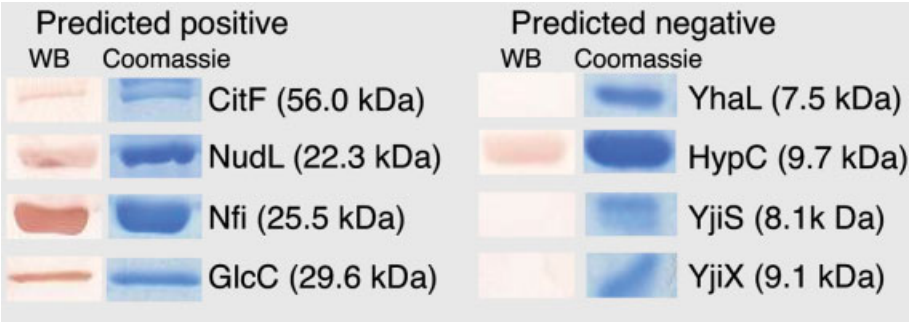


Figure 3. Experimental verification of NetPhosBac predicted phosphoproteins. Western blot of purified His-tagged proteins separated by SDS-PAGE. Left panel: Four high-scoring candidates that NetPhosBac predicts to be phosphorylated. Right panel: Four proteins predicted not to be phosphorylated. Both Western blot (WB) and Coomassie-stained slices are shown.

Table 3. Phosphorylation sites determined by MS analysis for the tested *E. coli* proteins

Gene ID	Protein name	Amino acid position	Modified sequence	NetPhosBac score
gi87082357	Nfi	S142	FEPLS**p**SEPGALAPLMDK	0.533
gi87082357	Nfi	S143	FEPLSS**p**EPGALAPLMDK	0.540
gi1789353	GlcC	S108	ALLEGES**p**AR	0.611
gi1789353	GlcC	S33	VGQPLPS**p**ER	0.552
gi1788115	NudL	S75	HAGQVAFPGGAVDDTDAS**p**AIAAALR	0.536
gi1788115	NudL	S5	S**p**LTLDDFLSR	*0.308*
gi1786832	CitF	S145	GLLAEPVQIHS**p**HGGR	0.895
gi1786832	CitF	S466	VVS**p**IEWLR	0.575
gi1786832	CitF	S133	GPLAEEIS**p**R	*0.413*
gi1786832	CitF	T310	ADFALGGITAT**p**MVDLHEK	*0.189*

Peptides identified by MS as phosphorylated (**p**) are shown with their respected NetPhosBac prediction scores. A positive prediction corresponds to a score above 0.5. False negative predictions are indicated in italics.

These results suggest that NetPhosBac is indeed capable of predicting phosphorylation of bacterial proteins and this is easily verifiable with a very accessible technique such as Western blot (Fig. 3). However, antiphosphoamino acid antibodies are not state of the art in phosphoprotein detection, since their binding depends on the sterical environment of the phosphorylated residue and they do not yield site-specific data. We therefore performed MS analysis of the four high-scoring candidates. In total, ten phosphopeptides were unambiguously identified at a site-specific level (Table 3, for MS spectra see Fig. S1 of Supporting Information). Out of the ten phosphorylation sites, seven were classified as true positives (NetPhosBac score >0.5) and three as false negatives. These experiments substantiate that NetPhosBac can reasonably be used to screen for potential substrates of protein kinases or distinguish a phosphorylated protein from a nonphosphorylated one.

4 Concluding remarks

Protein phosphorylation is a global regulatory mechanism, employed by all forms of life. Nevertheless, in the course of evolution, specific phylogenetic groups have evolved different types of kinases, which means that global phosphorylation predictors will always fall short in their prediction capacity when it comes to idiosyncratic taxa-specific kinases. NetPhosYeast and NetPhosBac illustrate the advantage of developing taxa-specific predictors when sufficient data become available. NetPhosBac demonstrated its competitive advantage on bacterial datasets and we expect it to further improve with the accumulation of new phosphorylation sites detected bacteria. One must remain careful as to its applicability to any bacterial system, since there are bacteria such as *Myxococcus xanthus* [43] and *Mycobacterium tuberculosis* [44] that contain numerous eukaryal-type kinases and in such cases predictors like the original NetPhos might give better results. Nevertheless, we are hopeful that NetPhosBac, trained on two model organisms representative of Gram-positive and Gram-negative bacteria, will provide a valuable tool to bacteriologists.

The authors would like to thank members of the Mann group for enlightening discussions. This work was supported by grants from the Danish Natural Science Research Council (FNU) and Interaction Proteome, contract number: LSHG-CT-2003-505520, European Commission FP6 Programme.

The authors have declared no conflict of interest.

5 References

[1] Pawson, T., Scott, J. D., Protein phosphorylation in signaling–50 years and counting. *Trends Biochem. Sci.* 2005, *30*, 286–290.

[2] Cohen, P., The role of protein phosphorylation in human health and disease. The Sir Hans Krebs Medal Lecture. *Eur. J. Biochem.* 2001, *268*, 5001–5010.

[3] Ptacek, J., Devgan, G., Michaud, G., Zhu, H. *et al.*, Global analysis of protein phosphorylation in yeast. *Nature* 2005, *438*, 679–684.

[4] Fabret, C., Feher, V. A., Hoch, J. A., Two-component signal transduction in *Bacillus subtilis*: How one organism sees its world. *J. Bacteriol.* 1999, *181*, 1975–1983.

[5] Hoch, J. A., Two-component and phosphorelay signal transduction. *Curr. Opin. Microbiol.* 2000, *3*, 165–170.

[6] Deutscher, J., Francke, C., Postma, P. W., How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol. Mol. Biol. Rev.* 2006, *70*, 939–1031.

[7] Cortay, J. C., Rieul, C., Duclos, B., Cozzone, A. J. *et al.*, Characterization of the phosphoproteins of *Escherichia coli* cells by electrophoretic analysis. *Eur. J. Biochem.* 1986, *159*, 227–237.

[8] Kennelly, P. J., Protein kinases and protein phosphatases in prokaryotes: A genomic perspective. *FEMS Microbiol. Lett.* 2002, *206*, 1–8.

[9] Bendt, A. K., Burkovski, A., Schaffer, S., Bott, M. *et al.*, Towards a phosphoproteome map of *Corynebacterium glutamicum*. *Proteomics* 2003, *3*, 1637–1646.

[10] Deutscher, J., Saier, M. H., Ser/Thr/Tyr protein phosphorylation in bacteria – for long time neglected, now well established. *J. Mol. Microbiol. Biotechnol.* 2005, *9*, 125–131.

[11] Grangeasse, C., Cozzone, A. J., Deutscher, J., Mijakovic, I. *et al.*, Tyrosine phosphorylation: An emerging regulatory device of bacterial physiology. *Trends Biochem. Sci.* 2007, *32*, 86–94.

[12] LaPorte, D. C., Chung, T., A single gene codes for the kinase and phosphatase which regulate isocitrate dehydrogenase. *J. Biol. Chem.* 1985, *260*, 15291–15297.

[13] Deutscher, J., Saier, M. H., ATP-dependent protein kinase-catalyzed phosphorylation of a seryl residue in HPr, a phosphate carrier protein of the phosphotransferase system in Streptococcus pyogenes. *Proc. Natl. Acad. Sci. USA* 1983, *80*, 6790–6794.

[14] Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S. *et al.*, Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 2004, *4*, 1633–1649.

[15] Huang, H. D., Lee, T. Y., Tzeng, S. W., Horng, J. T. *et al.*, KinasePhos: A web tool for identifying protein kinase-specific phosphorylation sites. *Nucleic Acids Res.* 2005, *33*, W226–W229.

[16] Xue, Y., Zhou, F., Zhu, M., Ahmed, K. *et al.*, GPS: A comprehensive www server for phosphorylation sites prediction. *Nucleic Acids Res.* 2005, *33*, W184–W187.

[17] Soufi, B., Jers, C., Hansen, M. E., Petranovic, D. *et al.*, Insights from site-specific phosphoproteomics in bacteria. *Biochim. Biophys. Acta* 2007, *7*, 299–307.

[18] Wurgler-Murphy, S. M., King, D. M., Kennelly, P. J., The Phosphorylation Site Database: A guide to the serine-, threonine-, and/or tyrosine-phosphorylated proteins in prokaryotic organisms. *Proteomics* 2004, *4*, 1562–1570.

[19] Macek, B., Mijakovic, I., Olsen, J. V., Gnad, F. *et al.*, The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. *Mol. Cell. Proteomics* 2007, *6*, 697–707.

[20] Macek, B., Gnad, F., Soufi, B., Kumar, C. *et al.*, Phosphoproteome analysis of *E. coli* reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Mol. Cell. Proteomics* 2008, *7*, 299–307.

[21] Li, W., Godzik, A., Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 2006, *22*, 1658–1659.

[22] Blom, N., Gammeltoft, S., Brunak, S., Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J. Mol. Biol.* 1999, *294*, 1351–1362.

[23] Ingrell, C. R., Miller, M. L., Jensen, O. N., Blom, N. *et al.*, NetPhosYeast: Prediction of protein phosphorylation sites in yeast. *Bioinformatics* 2007, *23*, 895–897.

[24] Iakoucheva, L. M., Radivojac, P., Brown, C. J., O'Connor, T. R. *et al.*, The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* 2004, *32*, 1037–1049.

[25] Levine, A., Vannier, F., Absalon, C., Kuhn, L. *et al.*, Analysis of the dynamic Bacillus subtilis Ser/Thr/Tyr phosphoproteome implicated in a wide variety of cellular processes. *Proteomics* 2006, *6*, 2157–2173.

[26] Eymann, C., Becher, D., Bernhardt, J., Gronau, K. *et al.*, Dynamics of protein phosphorylation on Ser/Thr/Tyr in Bacillus subtilis. *Proteomics* 2007, *7*, 3509–3526.

[27] Workman, C. T., Yin, Y., Corcoran, D. L., Ideker, T. *et al.*, enoLOGOS: A versatile web tool for energy normalized sequence logos. *Nucleic Acids Res.* 2005, *33*, W389–W392.

[28] Schneider, T. D., Stephens, R. M., Sequence logos: A new way to display consensus sequences. *Nucleic Acids Res.* 1990, *18*, 6097–6100.

[29] Diella, F., Cameron, S., Gemund, C., Linding, R. *et al.*, Phospho.ELM: A database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics* 2004, *5*, 79.

[30] Vacic, V., Iakoucheva, L. M., Radivojac, P., Two Sample Logo: A graphical representation of the differences between two sets of sequence alignments. *Bioinformatics* 2006, *22*, 1536–1537.

[31] Bendtsen, J. D., Nielsen, H., von Heijne, G., Brunak, S. *et al.*, Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 2004, *340*, 783–795.

[32] Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E. L. *et al.*, Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* 2001, *305*, 567–580.

[33] Amrein, K. E., Takacs, B., Stieger, M., Molnos, J. *et al.*, Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL. *Proc. Natl. Acad. Sci. USA* 1995, *92*, 1048–1052.

[34] Bork, P., Koonin, E. V., Protein sequence motifs. *Curr. Opin. Struct. Biol.* 1996, *6*, 366–376.

[35] Songyang, Z., Carraway, K. L., Eck, M. J., Harrison, S. C. *et al.*, Catalytic specificity of protein-tyrosine kinases is critical for selective signalling. *Nature* 1995, *373*, 536–539.

[36] Pawson, T., Regulation and targets of receptor tyrosine kinases. *Eur. J. Cancer* 2002, *38*, S3–S10.

[37] Krupa, A., Srinivasan, N., Diversity in domain architectures of Ser/Thr kinases and their homologues in prokaryotes. *BMC Genomics* 2005, *6*, 129.

[38] Kim, J. H., Lee, J., Oh, B., Kimm, K. *et al.*, Prediction of phosphorylation sites using SVMs. *Bioinformatics* 2004, *20*, 3179–3184.

[39] Hjerrild, M., Stensballe, A., Rasmussen, T. E., Kofoed, C. B. *et al.*, Identification of phosphorylation sites in protein kinase A substrates using artificial neural networks and mass spectrometry. *J. Proteome Res.* 2004, *3*, 426–433.

[40] Zhou, F. F., Xue, Y., Chen, G. L., Yao, X. *et al.*, GPS: A novel group-based phosphorylation predicting and scoring method. *Biochem. Biophys. Res. Commun.* 2004, *325*, 1443–1448.

[41] Xue, Y., Li, A., Wang, L., Feng, H. *et al.*, PPSP: Prediction of PK-specific phosphorylation site with Bayesian decision theory. *BMC Bioinformatics* 2006, *7*, 163.

[42] Manke, I. A., Nguyen, A., Lim, D., Stewart, M. Q. *et al.*, MAPKAP kinase-2 is a cell cycle checkpoint kinase that regulates the G2/M transition and S phase progression in response to UV irradiation. *Mol. Cell* 2005, *17*, 37–48.

[43] Munoz-Dorado, J., Inouye, S., Inouye, M., Eukaryotic-like protein serine/threonine kinases in *Myxococcus xanthus*, a developmental bacterium exhibiting social behavior. *J. Cell Biochem.* 1993, *51*, 29–33.

[44] Wehenkel, A., Bellinzoni, M., Grana, M., Duran, R. *et al.*, Mycobacterial Ser/Thr protein kinases and phosphatases: Physiological roles and therapeutic potential. *Biochim. Biophys. Acta* 2008, *1784*, 193–202.

Bacterial Protein-Tyrosine Kinases

Lei Shi[1], Ahasanul Kobir[1], Carsten Jers[2] and Ivan Mijakovic*,[1]

[1]*Micalis UMR 1319, AgroParisTech-INRA, Domaine de Vilvert, 78352 Jouy-en-Josas, France France;* [2]*Center for Systems Microbiology, Technical University of Denmark, DK-2800 Lyngby, Denmark*

Abstract: Bacteria and Eukarya share essentially the same family of protein-serine/threonine kinases, also known as the Hanks-type kinases. However, when it comes to protein-tyrosine phosphorylation, bacteria seem to have gone their own way. Bacterial protein-tyrosine kinases (BY-kinases) are bacterial enzymes that are unique in exploiting the ATP/GTP-binding Walker motif to catalyze phosphorylation of protein tyrosine residues. Characterized for the first time only a decade ago, BY-kinases have now come to the fore. Important regulatory roles have been linked with these enzymes, via their involvement in exopolysaccharide production, virulence, DNA metabolism, stress response and other key functions of the bacterial cell. BY-kinases act through autophosphorylation (mainly in exopolysaccharide production) and phosphorylation of other proteins, which have in most cases been shown to be activated by tyrosine phosphorylation. Protein-tyrosine phosphorylation in bacteria is particular with respect to very low occupancy of phosphorylation sites *in vivo*; this has represented a major challenge for detection techniques. Only the recent breakthroughs in gel-free high resolution mass spectrometry allowed the systematic detection of phosphorylated tyrosines by phosphoprotomics studies in bacteria. Other pioneering studies conducted in recent years, such as the first structures of BY-kinases and biochemical and phyiological studies of new BY-kinase substrates significantly furthered our understanding of these enzymes and highlighted their importance in bacterial physiology. Having no orthologues in Eukarya, BY-kinases are receiving a growing attention from the biomedical field, since they represent a particularly promising target for anti-bacterial drug design.

Keywords: Protein-tyrosine kinase / BY-kinase / bacteria / protein phosphorylation / phosphoproteome / signal transduction / cellular regulation

BY-KINASES IN BACTERIAL PHYSIOLOGY

The literature in the research field of post-translational modifications of proteins abounds with statements that elaborate modifications found in *Eukarya* do not exist in *Bacteria*. On more than one occasion, this simply reflected the fact that nobody had made a serious effort looking for them. Recent examples of "missing" modifications finally found in bacteria are those of protein glycosylation [1], acetylation [2] and methylation [2]. In much the same spirit, protein-tyrosine phosphorylation was considered non-existant in bacteria, until the pioneering studies appeared in the mid nineties [3,4], identifying phosphorylated tyrosine residues in bacterial proteins. These first studies reported members of a new family of proteins that autophosphorylated on tyrosine, which were later named BY-kinases (for bacterial tyrosine kinases) [5], but are also known as polysaccharide copolymerases [6] due to multiple roles they play in the cellular physiology. BY-kinase-encoding genes have since been identified in a majority of sequenced bacterial genomes, and are thus considered widespread, if not ubiquitous, in the bacterial kingdom [7]. BY-kinases account for most tyrosine phosphorylation events identified in bacteria so far, the exception being some unusual two-component systems that phosphorylate on tyrosine. A typical BY-kinase consists of an N-terminal transmembrane loop (of variable size), followed by a cytosolic catalytic domain that contains the ATP-binding site and a C-terminal tail of autophosphorylatable tyrosine residues [5]. A distinctive feature of these enzymes is that they use a structural motif known as the P-loop (Walker motifs A and B) to constitute their active site. This feature is shared by only a few bacterial protein-kinases [8], whereas in *Eukarya*, Walker motifs are found in many ATP/GTPases [9], but not in protein kinases. In *Firmicutes*, the canonic BY-kinase gene has been split in two, encoding separately the transmembrane domain and the cytosolic kinase, both of which maintain a tight functional interaction despite being separated to independent polypeptide chains (Fig. **1**). The cause and functional consequences of this separation remain unclear, although some new leads have appeared in recent studies and these will be discussed in the following chapter.

In the first phase of research on BY-kinases, since their discovery in 1996 and up to 2003, these enzymes were considered exclusively as autophosphorylating kinases. Initially, BY-kinase autophosphorylation has been linked to the functional role of these enzymes in exopolysaccharide production in several bacterial systems that were studied, since they are usually encoded by genes in the large operons involved in biosynthesis and export of sugar polymers [10]. One of the most exhaustively studied systems concerning BY-kinase implication in exopolysaccharide synthesis is the *Proteobacteria* model organism *Escherichia coli*. *E. coli* possesses two BY-kinases: Wzc encoded a gene in the operon which participates in the biosynthesis of polysaccharide polymers [11]

*Address correspondence to this author at the Micalis UMR 1319, AgroParisTech-INRA, F-78850 Thiverval-Grignon, FRANCE,
Tel: +33(0)1 30 81 45 40, Fax: +33(0)1 30 81 54 57;
E-mail: Ivan.Mijakovic@grignon.inra.fr

1570-1646/10 $55.00+.00

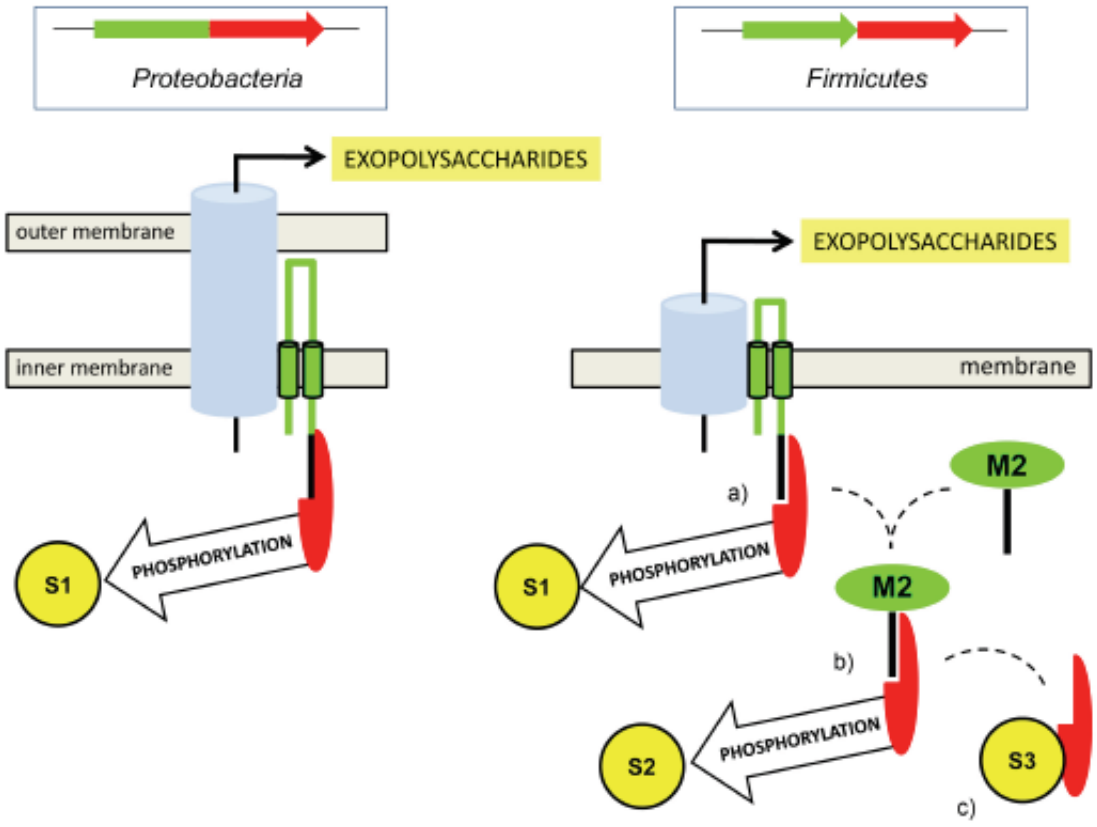


Fig. (1). Schematic representation of the two archetypal BY-kinase architectures. *Proteobacteria* have the two functional domains; transmembrane (green) and kinase (red) encoded by a single gene, and in *Firmicutes* these domains are independent proteins, encoded by separate genes. BY-kinases in both types of bacteria interact with the exopolysaccharide synthesis and translocation machinery (light blue), which is more complex in *Proteobacteria* where it has to traverse 2 membranes. The C-terminal end of the transmembrane domain that activates the kinase domain is represented in black. In *Proteobacteria*, this fragment is linked covalently to the kinase domain, and the active kinase domain is thus capable of phosphorylating the full set of its endogenous substrates (S1). In *Firmicutes*, the kinase is also activated by the C-terminal fragment of the transmembrane modulator, and can thus phosphorylate a set of physiological substrates (S1), as shown in case a). However, the kinase is free to dissociate from this modulator and meet other proteins that can accomplish this function (M2), and perhaps change its specificity towards another set of substrates (S2), as illustrated in case b). Finally, recent data indicate that the kinase can also co-localize with some of its substrates (S3), case c), although the physiological role of this co-localization is not clear at present.

and Etk, encoded by a gene in the G4C operon which is required for formation of group 4 capsule (G4C) polysaccharide [12]. It has been demonstrated that presence of Wzc and Etk is essential for synthesis of corresponding extracellular polysaccharides. More to the point, the auto-phosphorylation of tyrosine residues in the C-termini of these BY-kinases is the key feature in the assembly of capsular polysaccharides. Both the phosphorylated and the non-phosphorylated forms of Wzc are essential for polysaccharide polymer production [13,14], and the current hypothesis is that the cycling between phosphorylated and non-phosphorylated form of Wzc allows the polysaccharide polymer synthesis to proceed correctly. Accordingly, Wzc and Etk have been classified as polysaccharide co-polymerases. In their co-polymerase capacity, these BY-kinases have been shown to influence the amount of polysaccharides as well as the length of the produced polymer probably via an interaction with a polysaccharide polymerase Wzy [14-16]. An interesting observation from the evolutionary standpoint is that there are members of the polysaccharide co-polymerase family which contain the transmembrane domain, but have no cytosolic kinase domain [6]. A similar system has been described in *Firmicutes*, where the most extensively studied system with respect to exopolysaccharide production is probably that of *Streptococcus pneumoniae*. The autophosphorylating BY-kinase CpsD is essential for capsular polysaccharide biosynthesis in *S. pneumoniae*, where it regulates the amount of the synthesized capsular polysaccharide, which in turn affects the attachment of the bacteria to the host cells and contributes to the infection process [17,18]. Similar processes have been described in a number of other bacteria [5]. An interesting hypothesis has been put forward by Cuthbertson *et al.* [6] who argue that co-polymerases may simply influence the amount of produced polysaccharides by serving as molecular scaffolds for the other members of the translocon. The detailed mechanism of the functional interaction between polymerases and co-polymerases has not been entirely pinpointed, but considerable efforts are currently directed at this question.

After the initial link between BY-kinase autophosphorylation and exopolysaccharide production had been firmly

established, this field took a new direction in the year 2003, when several teams simultaneously published reports that BY-kinases can phosphorylate other endogenous proteins substrates on tyrosine residues [19-21]. In each reported case the substrate activity was regulated by BY-kinase dependent phosphorylation (Fig. **2**). *E. coli* Etk was found to phosphorylate RpoH (an alternative heat shock sigma factor) and an anti-sigma factor RseA and thus participate in triggering the heat shock response [19]. Its paralogue, Wzc [20], and a BY-kinase orthologue from *Bacillus subtilis*, PtkA [21], were found to phosphorylate UDP-glucose dehydrogenases in their respective organisms. The molecular mechanism of activation of UDP-glucose dehydrogenases by BY-kinase dependent tyrosine phosphorylation has been studied in molecular detail [22-24], and in the process, Etk from *E coli* has also been found to phosphorylate the cognate UDP-glucose dehydrogenases [23]. Although the mechanisms of UDP-glucose dehydrogenases activation are arguably not identical in *E. coli* [23] and *B. subtilis* [24], they both involve precise positioning of the phosphorylated tyrosine with respect to the bound substrate in the enzyme active site. Moreover, the investigation of the activation mechanism in *E. coli* has shed light on two parallel signal transduction pathways that intersect at the level of UDP-glucose dehydrogenase phosphorylation; one involving Wzc and the colanic acid synthesis and the other Etk and polymyxin resistance [23]. Examples of BY-kinases phosphorylating enzymes involved in the synthesis of sugar polymers have rapidly emerged in other bacterial systems. *S. thermophilus* phosphoglycosyltransferase EpsE is phosphorylated and activated by the cognate BY-kinase EpsD [25], *Staphylococcus aureus* BY-kinase Cap5B2 can phosphorylate and activate an endogenous UDP-acetyl-mannosamine dehydrogenase Cap5O [26], whereas tyrosine-phosphorylation of *Klebsiella pneumoniae* undecaprenolphosphate glycosyltransferase WcaJ has been reported as necessary for capsular polysaccharide synthesis [27]. The emergent phosphorylation networks based on BY-kinases and their substrates hold promise of significant complexity. So far only one case of a single substrate phosphorylated by several kinases has been reported: *E. coli* Ugd phosphorylated by Wzc and Etk [23]. However, the examples of one kinase phosphorylating several substrates are plenty: *E. coli* Etk phosphorylates RpoH, RseA [19] and Ugd [23], and *B. subtilis* PtkA phosphorylates UDP-glucose dehydrogenases Ugd and TuaD [21], as well as single-stranded DNA-binding proteins SsbA and SsbB [28]. Phosphorylation of single-stranded DNA-binding proteins by BY-kinases, which results in increased affinity for DNA, hinted at an interesting possibility that BY-kinases migh in fact be extensively involved in regulating the DNA metabolism. *B. subtilis* cells devoid of PtkA exhibited a complex pleiotropic phenotype with defects in the cell cycle, initiation of the DNA replication and chromosome distribution [29]. In addition to SsbA and SsbB, PtkA phosphorylates and regulates the activity of other proteins involved in *B. subtilis* single-standed DNA metabolism, which may account for the complex phenotype of the Δ*ptkA* strain (Jers & Mijakovic, unpublished results). In *E. coli*, the BY-kinase Wzc was found to phosphorylate integrase proteins (Int) of coliphages HK022 and λ. Overexpression of Wzc in tyrosine-phosphatase deficient background resulted in a significantly reduced lysogenization, indicating that phosphorylation of Int down-regulates its activity [30].

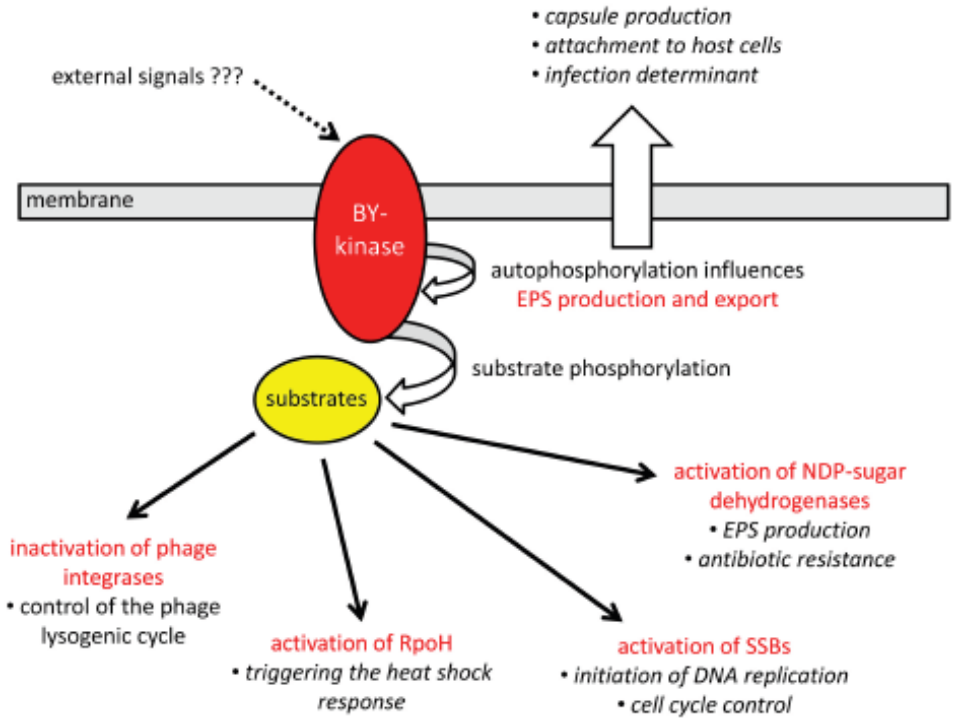


Fig. (2). Schematic representation of the physiological and regulatory roles of BY-kinases. BY-kinase autophosphorylation controls the involvement of these enzymes in the synthesis and translocation of sugar polymers. In addition, BY-kinases phosphorylate a number of intracellular substrates. Only well-studied cases are presented here, with clearly documented consequences for bacterial physiology.

Following the finding that heat shock response in *E. coli* is modulated by the activity of the BY-kinase Etk [19], a link between tyrosine phosphorylation and heat shock was also published for *B. subtilis*. A protein with no homology to canonical BY-kinases, McsB, was found to phosphorylate CtsR, repressor of the heat-shock genes in the presence of a modulator protein McsA [31]. By consequence, the tyrosine-phosphorylated CtsR was reported to release its target DNA. McsB was also found to act as a regulated adaptor protein for ClpCP, the protease complex responsible for the degradation of CtsR [32]. The active, phosphorylated kinase McsB interacts with both CtsR and ClpC, and targets CtsR to the ClpCP complex. Finally, in 2009 McsB seems to have been taken off the list of tyrosine kinases, and its status revised to that of an arginine kinase [33]. It would be interesting to see whether this kinase has a dual specificity, or the discrepancies between the two reports were due to different detection methods. This by no means diminished the importance of previous findings that McsB functionally interacts with CtsR and ClpCP. At the same time, this new finding opened a new chapter of bacterial protein phosphorylation at arginine residues.

PROTEOBACTERIA VS. *FIRMICUTES*: DIFFERENT STRUCTURES FOR DIFFERENT REGULATORY STRATEGIES?

Important advances in the field of bacterial protein-tyrosine kinases were made recently in terms of structural studies. The Walker motifs found in BY-kinases, and the overall sequence homology, had suggested that these proteins might structurally resemble bacterial ATPases such as MinD and Soj [5]. The first structural insights came from the low resolution structure (14 Å) of *E. coli* BY-kinase Wzc in complex with the capsular polysaccharide translocon Wza [34]. It revealed a Wzc tetramer that oligomerises via the periplasmic domains, and with its cytosolic kinase domains freely protruding into the cytosol (not interacting with each other). The functional insights derived from this structure suggested that Wzc regulates export by triggering an open/active conformation of the translocase Wza. More recently, two high resolution structures of BY-kinases were published: Etk from *E. coli* [35] representing the *Proteobacteria*-type BY-kinase architecture, and CapA/B from *S. aureus* [36] representing the *Firmicutes*-type BY-kinase architecture. These structures provided insights into the functional role of several conserved features. In a previous biochemical study it was demonstrated that Etk homologue Wzc is activated through a two-step mechanism involving a conserved tyrosine in the active site and the C-terminal tyrosine cluster [37]. The structure data combined with biochemical and *in silico* analysis revealed a novel mechanism by which the first activation steps proceeds, showing that the conserved tyrosine side chain points into the active site thereby blocking activity. Upon autophosphorylation, the negatively charged phosphotyrosine rotates out of the active site and is stabilized by interaction with an arginine residue [35,38]. In *Firmicutes*, where the kinase domain is separated in a distinct polypeptide chain, it has been shown that the interaction with the remaining transmembrane domain is necessary for the kinase activity [21,39]. The structure of *S. aureus* CapB demonstrated that the activation proceeds, at least in part, by the fact that the C-terminal part of the transmembrane modulator (protruding into the cytosol) stabilizes ATP binding by hydrophobic sandwich interaction between the adenine ring and a phenylalanine residue of the modulator protein [36]. In case of CapB, the non-phosphorylated kinase monomers, as opposed to the case of Wzc, associated in an octameric ring structure, anchored to the membrane via interaction with the transmembrane modulator CapA. Notably, one of the tyrosine residues in the C-terminal tyrosine cluster was found bound to the active site of the neighbouring subunit, suggesting an inter-molecular phosphorylation mechanism. The fact that no tyrosine residue in the cluster is preferentially phosphorylated reflects a high degree of flexibility in that region. An important conclusion from this study was that autophosphorylation of CapB is expected to induce dissociation of the ring structure, by a conformational change that disrupts the amino acid contacts at the interaction interface, while maintaining the interaction between individual CapB-CapA couples. The two BY-kinase structures also provided a rationale for the substrate specificity of Tyr vs Ser/Thr kinases and the difference between BY-kinases and the structurally highly similar ATPases such as MinD [40].

The resolved structures of BY-kinases have rekindled interest in a decade-old question, why are there distinct *Proteobacteria*- and *Firmicutes*-type architectures; one with a single-polypeptide chain, and the other with two functional domains (transmembrane and kinase) split in separate proteins (Fig. **1**). BY-kinases in *Proteobacteria* have a large extracellular (periplasmic) loop that allows them to interact with the polysaccharide translocation machinery [6], and this loop is significantly shorter in *Firmicutes* which lack the outer membrane and all corresponding structures. However, both types of BY-kinases are perfectly capable of phosphorylating various intra-cellular protein substrates, so this particular activity (substrate phosphorylation and recognition of multiple substrates) does not seem to be linked with any one of the two particular architectures. The structural study on CapB indicated that upon octamer dissociation, each individual BY-kinase molecule remains in contact with its transmembrane modulator CapA, and were it ever to separate from it, it would become effectively inactive as a kinase, due to destabilization of the ATP-binding site [36]. And yet the fact remains, BY-kinases in *Firmicutes* have evolved with an option to dissociate from their transmembrane modulators, which makes it rather difficult to imagine that they never do so. Our group has recently performed localization studies on the BY-kinase PtkA in *B. subtilis* using fluorescent protein fusions, and we were able to demonstrate that PtkA co-localizes with its transmembrane modulator TkmA under certain growth conditions, but we were equally able to identify growth conditions where PtkA dissociates from TkmA and becomes prevalently cytosolic (Jers & Mijakovic, unpublished results). In some cases PtkA also co-localized with some of its newly identified proteins substrates. We further demonstrated that PtkA can interact with other cytosolic proteins that are not its substrates but can modulate its kinase activity (Shi, Noirot-Gros & Mijakovic, unpublished results). This incited us to speculate that BY-kinases in *Firmicutes* might have a separated kinase domain in order to make it accessible to other modulators, and therefore participate in signaling pathways other than the principal one that starts at

the cell membrane, with the classical BY-kinase transmembrane domain. Arguments to support our hypothesis are admittedly not abundant at the moment, but we are confident that further studies in this direction will bring about another important turning point in the story of BY-kinases.

DETECTING PHOSPHO-TYROSINES: DEVELOPMENTS IN MASS SPECTROMETRY PHOSPHOPROTEOMICS

In *Eukarya*, protein-tyrosine phosphorylation only accounts for approximately 0.05% of the total cellular protein phosphorylation, even though almost one fifth of the cellular kinome encodes protein-tyrosine kinases [41]. Despite the average scarcity in terms of occupancy of phosphorylation sites, protein-tyrosine phosphorylation plays key roles in eukaryal signal transduction, cell differentiation and growth. In *Bacteria*, occupancy of protein-tyrosine phosphorylation sites is also very low, to the extent that this type of phosphorylation has not been identified at all by traditional proteomics approaches using 2D-gels and low-resolution mass spectrometry [42]. Only recently, the advent of gel-free analysis coupled to high-resolution mass spectrometry enabled us to systematically detect phospho-tyrosines in bacterial phosphoproteomes [27, 43-45]. In terms of the number of identified phosphorylation sites, phospho-tyrosines account for 3-10% of published bacterial phosphoproteomes and this is a considerable over-representation compared to the eukaryal systems. Nevertheless, in terms of actual site occupancy tyrosine phosphorylation remains very low, and the signals of tyrosine-phosphorylated peptides during mass spectrometry analysis are as a rule difficult to detect due to ion suppression effects from a comparatively high background of non-phosphorylated peptides. Therefore, enrichment techniques for phosphotyrosine-containing peptides preceding mass spectrometry analysis represent a promising area for improvement. In all phosphoproteomic studies performed in bacteria so far, phosphopeptides were enriched by methods not distinguishing between various phosphoesters (phospho-serine, -threonine and -tyrosine), such as using strong cationic exchange, titanium-oxide or immobilized metal affinity chromatography. Anti-phosphoamino acid antibodies against phospho-serine/threonine and to lesser extent against phospho-tyrosine are infamous for somewhat relaxed specificity, due to the fact that antibodies recognize a stretch of 10-11 amino acids and a single phosphoamino acid is not an ideal epitope. While this may represent a challenge for specific detection by Western blots, in terms of phosphopeptide enrichment from crude protein extracts antibodies can be a useful tool. In studies performed on eukaryal systems, the 4G10 and the PY100 antibodies were found to be the most efficient for the immune-affinity enrichment of phosphotyrosine proteins and peptides, and a combined use of the two antibodies was recommended to obtain the best result [46]. It seems therefore that a viable route towards global tyrosine phosphorylation studies in bacteria might imply antibody based enrichment prior to mass spectrometry analysis. In terms of cellular signaling, protein-tyrosine phosphorylation is a dynamic event and quantitative and time resolved analytical approaches are prerequisite to understanding its physiological role; even more so bearing in mind the short generation time of bacteria. In quantitative phosphoproteomics, stable isotope labeling of cell cultures with amino acids (SILAC) provides reliable means of relative quantification of phosphotyrosines [47]. However, SILAC-based quantification imposes some limitations in terms of choice of growth medium. Alternative approaches rely on labeling after the growth phase, thus not influencing the growth conditions, such as for example the dimethylation of peptides with stable isotopes [48]. In parallel to mass spectrometry-based phosphoproteomics, unrelated approaches are being developed to study protein-tyrosine phosphorylation. They may lack the broadness of the global phosphoproteomic approaches, but may have other advantages in terms of time resolved detection of specific phosphorylation events. For example, electrochemical biosensors have been developed to monitor specific tyrosine kinase activities (and screening of inhibitors) *in vitro* [49]. In this particular study, the authors have used specific biotinylated substrate peptides immobilized on a streptavidin-coated indium-tin oxide electrode surface. The phosphorylation is detected using an ATP analogue adenosine gamma-thio triphosphate. After kinase-catalysed thio-phosphorylation of the immobilized peptides, the electrode surface was exposed to gold nanoparticles and the signals were thereafter detected by voltametry, using gold-chloride electrochemistry. This method seems particularly promising for *in vitro* screening of specific tyrosine-kinase inhibitors, and other developments in nanotechnology biosensors may hold a promise for applications allowing *in vivo* studies of signal transduction in microorganisms.

PERSPECTIVES

Site-specific bacterial phosphoproteomes are being published at an increasing pace [27, 43-45], and with the high accuracy mass spectrometry becoming more accessible we may reasonably expect to identify large numbers of BY-kinase substrates in various bacterial organisms. At present, bioinformatic predictors of protein phosphorylation sites are available. Neural-network algorithms, such as NetPhosBac [50], trained with available datasets on experimentally identified phosphorylation sites, can predict serine/threonine phosphorylation based on sequence homology and conservation of sites. Alternative predictors are the ones based on local structural disorder, such as DISPHOS [51]. At present, the number of identified sites of protein-tyrosine phosphorylation in bacterial proteins is not sufficient to envisage the construction of homology-based predictors such as NetPhosBac, but approaches based on local structural disorder seem more promising. In particular, weak conservation of bacterial phosphoproteomes on the phosphorylation site level is an argument in favour of developing clade-specific predictors, rather than global bacterial predictors. Upon examination of published tyrosine phosphorylation sites assigned to a specific kinase, as for example Ugd, TuaD, SsbA and SsbB proteins phosphorylated by PtkA in *B. subtilis* [21,28], no clear kinase recognition motif has emerged in the immediate surroundings of the phosphorylated residues. It is therefore plausible to assume that BY-kinases, much like their cognate phosphotyrosine-protein phosphatases [52], recognize structural motifs distant from the phosphorylated(able) residue on which they act. One big remaining question concerning the regulatory role of BY-kinases is the regulation of their kinase activity. BY-kinases described in recent structural

studies offer a snapshot of a constitutively active kinase, but we cannot be certain that this is indeed the case *in vivo*. The presence of large extracellular loops, by analogy to eukaryal systems, suggests the existence of extracellular signals that might trigger kinase activity. However, in *Proteobacteria* these extracellular loops are heavily involved in interactions with polysaccharide translocation machinery, and it is questionable whether they may form a real signal acceptor site. In bacteria in general, our knowledge on extracellular signals that trigger the activity of cytosolic serine/threonine and tyrosine kinases is very limited [53], but this may simply reflect our lack of investigative approaches concerning this part of the signaling cascade. The amount of data gathered on BY-kinases is at present not extensive enough to allow in-depth modeling of tyrosine-phosphorylation-based regulatory networks. However, the available elements on specific substrate activation should soon allow the construction of simple Boolean-type models, integrating all known BY-kinase interactants. Since tyrosine phosphorylation is a dynamic event; it may be justified to invest some efforts in quantitative, time-resolved analysis of BY-kinase cellular functions, and create dynamic models describing the behavior of these regulatory loops. These efforts may be particularly well justified by the relatively high connectivity of BY-kinases in Firmicutes, with a large number of substrates and alternative modulators.

Recent reviews concerning BY-kinases and bacterial protein phosphorylation in general [54,55] point out the promise they hold in terms of strategies for battling infective diseases caused by bacterial pathogens. The main advantage of BY-kinases in this respect is the fact that they have no eukaryal homologues, so their specific inhibitors are less likely to affect the host than for example the inhibitors of Hanks type serine/threonine kinases, which are present both in *Bacteria* and *Eukarya*. However, as discussed above, BY-kinases seem to be quite promiscuous in terms of choosing their substrates. Along these lines, Lin *et al.* [27] justifiably argue that site-specific knowledge concerning the substrates regulated by BY-kinases might provide more specific targets for inhibitory drugs. There is nevertheless an agreement in the field that inhibiting bacterial tyrosine phosphorylation is a promising venue in developing new therapeutics, and considerable efforts are already committed to this approach.

ACKNOWLEDGEMENTS

This work was supported by grants from the Lundbeckfonden and the Institut National de Recherche Agronomique (INRA) to IM and a PhD stipend from the Technical University of Denmark (DTU) to CJ.

References

[1] Abu-Qarn, M.; Eichler, J. and Sharon N. Not just for Eukarya anymore: protein glycosylation in Bacteria and Archaea. *Curr. Opin. Struct. Biol*., **2008**, *18*, 544-550.

[2] Lauber, M.A.; Running, W.E. and Reilly, J.P. *B. subtilis* ribosomal proteins: structural homology and post-translational modifications. *J. Proteome Res*., **2009**, *8*, 4193-4206.

[3] Duclos, B.; Grangeasse, C.; Vaganay, E.; Riberty, M. and Cozzone, A.J. Autophosphorylation of a bacterial protein at tyrosine. *J. Mol. Biol.*, **1996**, *259*, 891-895.

[4] Grangeasse, C.; Doublet, P.; Vaganay, E.; Vincent, C.; Deléage, G.; Duclos, B. and Cozzone, A.J. Characterization of a bacterial gene encoding an autophosphorylating protein tyrosine kinase. *Gene*, **1997**, *204*, 259-265.

[5] Grangeasse, C.; Cozzone, A.J.; Deutscher, J. and Mijakovic, I. Tyrosine phosphorylation: an emerging regulatory device of bacterial physiology. *Trends Biochem. Sci.*, **2007**, *32*, 86-94.

[6] Cuthbertson, L.; Mainprize, I.L.; Naismith, J.H. and Whitfield, C. Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in Gram-negative bacteria. *Microbiol. Mol. Biol. Rev*., **2009**, *73*, 155-177.

[7] Jadeau, F.; Bechet, E.; Cozzone, A.J.; Deléage, G.; Grangeasse, C. and Combet, C. Identification of the idiosyncratic bacterial protein tyrosine kinase (BY-kinase) family signature. *Bioinformatics*, **2008**, *24*, 2427-2430.

[8] Mijakovic, I.; Poncet, S.; Galinier, A.; Monedero, V.; Fieulaine; S., Janin, J.; Nessler, S.; Marquez, H.-A.; Scheffzek, K.; Hasenbein, S.; Hengstenberg, W. and Deutscher, J. Pyrophosphate-producing protein dephosphorylation by HPr kinase/phosphorylase: a relic of early life? *Proc. Natl. Acad. Sci. USA*, **2002**, *99*, 13442-13347.

[9] Patel, S. and Latterich, M. The AAA team: related ATPases with diverse functions. *Trends Cell. Biol.*, **1998**, *8*, 65-71.

[10] Whitfied, C. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli. Annu. Rev. Biochem*., **2006**, *75*, 39-68.

[11] Vincent, C.; Doublet, P.; Grangeasse, C.; Vaganay, E.; Cozzone, A.J. and Duclos, B. Cells of *Escherichia coli* contain a protein-tyrosine kinase, Wzc, and a phosphotyrosine-protein phosphatase, Wzb. *J. Bacteriol.*, **1999**, *181*, 3472-3477.

[12] Peleg, A.; Shifrin, Y.; Ilan, O.; Nadler-Yona, C.; Nov, S.; Koby, S.; Baruch, K.; Altuvia, S.; Elgrably-Weiss, M.; Abe, C.M.; Knutton, S.; Saper, M.A. and Rosenshine, I. Identification of an *Escherichia coli* operon required for formation of the O-antigen capsule. *J. Bacteriol.*, **2005**, *187*, 5259-5266.

[13] Paiment, A.; Hocking, J. and Whitfield, C. Impact of phosphorylation of specific residues in the tyrosine autokinase, Wzc, on its activity in assembly of group 1 capsules in *Escherichia coli*. *J. Bacteriol.*, **2002**, *184*, 6437-6447.

[14] Obadia, B.; Lacour, S.; Doublet, P.; Baubichon-Cortay, H.; Cozzone, A.J. and Grangeasse, C. Influence of tyrosine-kinase Wzc activity on colanic acid production in *Escherichia coli* K12 cells. *J. Mol. Biol.*, **2007**, *367*, 42-53.

[15] Tocilj, A.; Munger, C.; Proteau, A.; Morona, R.; Purins, L.; Ajamian, E.; Wagner, J.; Papadopoulos, M.; Van Den Bosch, L.; Rubinstein, J.L.; Féthière, J.; Matte, A. and Cygler, M. Bacterial polysaccharide co-polymerases share a common framework for control of polymer length. *Nat. Struct. Mol. Biol.*, **2008**, *15*, 130-138.

[16] Whitfield, C. and Larue, K. Stop and go: regulation of chain length in the biosynthesis of bacterial polysaccharides. *Nat. Struct. Mol. Biol.*, **2008**, *15*, 121-123.

[17] Morona, J.K.; Morona, R. and Paton, J.C. Attachment of capsular polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for invasive disease. *Proc. Natl. Acad. Sci. USA*, **2006**, *103*, 8505-8510.

[18] Morona, J.K.; Miller, D.C.; Morona, R. and Paton, J.C. The effect that mutations in the conserved capsular polysaccharide biosynthesis genes cpsA, cpsB, and cpsD have on virulence of *Streptococcus pneumoniae*. *J. Infect. Dis.*, **2004**, *189*, 1905-1913.

[19] Klein, G.; Dartigalongue, C. and Raina, S. Phosphorylation-mediated regulation of heat shock response in *Escherichia coli*. *Mol. Microbiol*., **2003**, *48*, 269-285.

[20] Grangeasse, C.; Obadia, B.; Mijakovic, I.; Deutscher, J.; Cozzone, A.J. and Doublet, P. Autophosphorylation of the *Escherichia coli* protein kinase Wzc regulates tyrosine phosphorylation of Ugd, a UDP-glucose dehydrogenase. *J. Biol. Chem*., **2003**, *278*, 39323-39329.

[21] Mijakovic, I.; Poncet, S.; Boël, G.; Mazé, A.; Gillet, S.; Jamet, E.; Decottignies, P.; Grangeasse, C.; Doublet, P.; Le Maréchal, P. and Deutscher, J. Transmembrane modulator-dependent bacterial tyrosine kinase activates UDP-glucose dehydrogenases. *EMBO J*., **2003**, *22*, 4709-4718.

[22] Mijakovic, I.; Petranovic, D. and Deutscher, J. How tyrosine phosphorylation affects the UDP-glucose dehydrogenase activity of *Bacillus subtilis* YwqF. *J. Mol. Microbiol. Biotechnol*., **2004**, *8*, 19-25.

[23] Lacour, S.; Bechet, E.; Cozzone, A.J.; Mijakovic, I. and Grangeasse, C. Tyrosine phosphorylation of the UDP-glucose de-

hydrogenase of *Escherichia coli* is at the crossroads of colanic acid synthesis and polymyxin resistance. *PLoS One*, **2008**, *3*, e3053.

[24] Petranovic, D.; Grangeasse, C.; Macek, B.; Abdillatef, M.; Gueguen-Chaignon, V.; Nessler, S.; Deutscher, J. and Mijakovic, I. Activation of *Bacillus subtilis* Ugd by the BY-kinase PtkA proceeds via phosphorylation of its residue tyrosine 70. *J. Mol. Microbiol. Biotechnol.*, **2009**, *17*, 83-89.

[25] Minic, Z.; Marie, C.; Delorme, C.; Faurie, J.M.; Mercier, G.; Ehrlich, D. and Renault, P. Control of EpsE, the phosphoglycosyltransferase initiating exopolysaccharide synthesis in *Streptococcus thermophilus*, by EpsD tyrosine kinase. *J. Bacteriol.*, **2007**, *189*, 1351-1357.

[26] Soulat, D.; Grangeasse, C.; Vaganay, E.; Cozzone, A.J. and Duclos, B. UDP-acetyl-mannosamine dehydrogenase is an endogenous protein substrate of *Staphylococcus aureus* protein-tyrosine kinase activity. *J. Mol. Microbiol. Biotechnol.*, **2007**, *13*, 45-54.

[27] Lin, M.H.; Hsu, T.L.; Lin, S.Y.; Pan, Y.J.; Jan, J.T.; Wang, J.T.; Khoo, K.H. and Wu, S.H. Phosphoproteomics of *Klebsiella pneumonia* NTUH-K2044 reveals a tight link between tyrosine phosphorylation and virulence. *Mol. Cell. Proteom.*, **2009**, in press, PMID: 19696081.

[28] Mijakovic, I.; Petranovic, D.; Macek, B.; Cepo, T.; Mann, M.; Davies, J.; Jensen, P.R.; and Vujaklija, D. Bacterial single-stranded DNA-binding proteins are phosphorylated on tyrosine. *Nucleic Acids Res.*, **2006**, *34*, 1588-1596.

[29] Petranovic, D.; Michelsen, O.; Zahradka, K.; Silva, C.; Petranovic, M.; Jensen, P.R. and Mijakovic, I. *Bacillus subtilis* strain deficient for the protein-tyrosine kinase PtkA exhibits impaired DNA replication. *Mol. Microbiol.*, **2007**, *63*, 1797-1805.

[30] Kolot, M.; Gorovits, R.; Silberstein, N.; Fichtman, B. and Yagil, E. Phosphorylation of the integrase protein of coliphage HK022. *Virology*, **2008**, *375*, 383-390.

[31] Kirstein, J.; Zühlke, D.; Gerth, U.; Turgay, K. and Hecker, M. A tyrosine kinase and its activator control the activity of the CtsR heat shock repressor in *B. subtilis*. *EMBO J.*, **2005**, *24*, 3435-3445.

[32] Kirstein, J.; Dougan, D.A.; Gerth, U.; Hecker, M. and Turgay, K. The tyrosine kinase McsB is a regulated adaptor protein for ClpCP. *EMBO J.*, **2007**, *26*, 2061-2070.

[33] Fuhrmann, J.; Schmidt, A.; Spiess, S.; Lehner, A.; Turgay, K.; Mechtler, K.; Charpentier, E. and Clausen, T. McsB is a protein arginine kinase that phosphorylates and inhibits the heat-shock regulator CtsR. *Science*, **2009**, *324*, 1323-1327.

[34] Collins, R.F.; Beis, K.; Dong, C.; Botting, C.H.; McDonnell, C.; Ford, R.C.; Clarke, B.R.; Whitfield, C. and Naismith, J.H. The 3D structure of a periplasm-spanning platform required for assembly of group 1 capsular polysaccharides in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **2007**, *104*, 2390-2395.

[35] Lee, D.C.; Zheng, J.; She, Y.M. and Jia, Z. Structure of *Escherichia coli* tyrosine kinase Etk reveals a novel activation mechanism. *EMBO J.*, **2008**, *27*, 1758-1766.

[36] Olivares-Illana, V.; Meyer, P.; Bechet, E.; Gueguen-Chaignon, V.; Soulat, D.; Lazereg-Riquier, S.; Mijakovic, I.; Deutscher, J.; Cozzone, A.J.; Laprévote, O.; Morera, S.; Grangeasse, C. and Nessler, S. Structural basis for the regulation mechanism of the tyrosine kinase CapB from *Staphylococcus aureus*. *PLoS Biol.*, **2008**, *6*, e143.

[37] Grangeasse, C.; Doublet, P. and Cozzone, A.J. Tyrosine phosphorylation of protein kinase Wzc from *Escherichia coli* K12 occurs through a two-step process. *J. Biol. Chem.*, **2002**, *277*, 7127-7135.

[38] Lu, T.; Tan, H.; Lee, D.; Chen, G. and Jia, Z. New insights into the activation of *Escherichia coli* tyrosine kinase revealed by molecular dynamics simulation and biochemical analysis. *Biochemistry*, **2009**, *48*, 7986-7995.

[39] Soulat, D.; Jault, J.M.; Duclos, B.; Geourjon, C.: Cozzone, A.J.; and Grangeasse, C. *Staphylococcus aureus* operates protein-tyrosine phosphorylation through a specific mechanism. *J. Biol. Chem*., **2006**, *281*, 14048-14056.

[40] Lee, D.C. and Jia, Z. Emerging structural insights into bacterial tyrosine kinases. *Trends Biochem. Sci.*, **2009**, *34*, 351-357.

[41] Blume-Jensen. P. and Hunter, T**.** Oncogenic kinase signaling. *Nature*, **2001**, *411*, 355-365.

[42] Bendt, A.K.; Burkovski, A.; Schaffer, S.; Bott, M.; Farwick, M. and Hermann, T. Towards a phosphoproteome map of *Corynebacterium glutamicum*. *Proteomics*, **2003**, *3*, 1637-1646.

[43] Macek, B.; Mijakovic, I.; Olsen, J.V.; Gnad, F.; Kumar, C.; Jensen, P.R. and Mann, M. The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. *Mol. Cell. Proteom.*, **2007**, *6*, 697-707.

[44] Macek, B.; Gnad, F.; Soufi, B.; Kumar, C.; Olsen, J.V.; Mijakovic, I. and Mann, M. Phosphoproteome analysis of *E. coli* reveals evolutionary conservation of bacterial Ser/Thr/Tyr phosphorylation. *Mol. Cell. Proteom.*, **2008**, *7*, 299-307.

[45] Soufi, B.; Gnad, F.; Jensen, P.R.; Petranovic, D.; Mann, M.; Mijakovic, I. and Macek, B. The Ser/Thr/Tyr phosphoproteome of *Lactococcus lactis* IL1403 reveals multiply phosphorylated proteins. *Proteomics*, **2008**, *8*, 3486-3493.

[46] Lind, S.B.; Molin, M.; Savitski, M.M.; Emilsson, L.; Åström, J.; Hedberg, L., Adams, C.; Nielsen, M.L., Engström, A.; Elfineh, L.; Andersson, E.; Zubarev, R.A. and Pettersson, U. Immunoaffinity enrichments followed by mass spectrometric detection for studying global protein tyrosine phosphorylation. *J. Prot. Res.*, **2008**, *7*, 2897-2910.

[47] Bantscheff, M.; Schirle, M.; Sweetman, G.; Rick, J. and Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.*, **2007**, *389*, 1017-1031.

[48] Boersema, P.J.; Aye, T.T.; van Veen, T.A.; Heck, A.J. and Mohammed, S. Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. *Proteomics*, **2008**, *8*, 4624-4632.

[49] Kerman, K. and Kraatz, H.B. Electrochemical detection of protein tyrosine kinase-catalysed phosphorylation using gold nanoparticles. *Biosens. Bioelectron.*, **2009**, *24*, 1484-1489.

[50] Miller, M.L.; Soufi, B.; Jers, C.; Blom, N.; Macek, B. and Mijakovic, I. NetPhosBac - a predictor for Ser/Thr phosphorylation sites in bacterial proteins. *Proteomics*, **2009**, 9, 116-125.

[51] Iakoucheva, L.M.; Radivojac, P.; Brown, C.J.; O'Connor, T.R.; Sikes, J.G.; Obradovic, Z. and , Dunker, A.K. Intrinsic disorder and protein phosphorylation. *Nucleic Acids Res.*, **2004**, *32*, 1037-1049.

[52] Mijakovic, I.; Musumeci, L.; Tautz, L.; Petranovic, D.; Edwards, R.A.; Jensen, P.R.; Mustelin, T.; Deutscher, J. and Bottini N. *In vitro* characterization of the *Bacillus subtilis* protein tyrosine phosphatase YwqE. *J. Bacteriol.*, **2005**, 187, 3384-3390.

[53] Shah, I.M.; Laaberki, M.H.; Popham, D.L. and Dworkin, J. A eukaryotic-like Ser/Thr kinase signals bacteria to exit dormancy in response to peptidoglycan fragments. *Cell*, **2008**, *135*, 486-496.

[54] Grangeasse, C.; Terreux, R. and Nessler, S. Bacterial tyrosine-kinases: structure-function analysis and therapeutic potential. *Biochim. Biophys. Acta*, **2009**, in press, doi: 10.1016/j.bbapap.2009.08.018

[55] Jers, C.; Soufi, B.; Grangeasse, C.; Deutscher, J. and Mijakovic, I. Phosphoproteomics in bacteria: towards a systemic understanding of bacterial phosphorylation networks. *Exp. Rev. Proteom.*, **2008**, *5*, 619-627.

Received: November 25, 2009 Revised: May 11, 2010 Accepted: May 14, 2010

Joint author statement

If a thesis contains articles made in collaboration with other researchers, a joint author statement about the PhD-student's part of the article shall be made by each of the co-authors, cf. article 12, section 4 of the Ministerial Order No. 18 February 2008 about the PhD degree

Title of the article:	*Bacillus subtilis* BY-kinase PtkA controls enzyme activity and localisation of its protein substrates.
Author(s):	Jers, C, Pedersen, M.M., Paspaliari, D.K., Schütz, W., Johnsson, C., Soufi, B., Macek, B., Jensen, P.R. and Mijakovic, I.
Journal:	Molecular microbiology (submitted)
PhD-student:	Carsten Jers
CPR-no.:	180179-2835
Signature of the PhD-student:	
Date:	10 May 2010

Co-author:	Malene Mejer Pedersen	Signature:	
Co-author:	Dafni Katerina Paspaliari	Signature:	
Co-author:	Wolfgang Schütz	Signature:	
Co-author:	Christina Johnsson	Signature:	
Co-author:	Boumediene Soufi	Signature:	
Co-author:	Boris Macek	Signature:	
Co-author:	Peter Ruhdal Jensen	Signature:	
Co-author:	Ivan Mijakovic	Signature:	Prof Dr Ivan MIJAKOVIĆ

Description of each author's contribution to the above-mentioned article:

Experiments were planned by IM and CJe. *in vitro* assays were done by DKP and CJe. Localisation experiments were done by MMP and CJe. Mass spectrometry analysis was done by WS and BM who also wrote the materials and methods part dealing with mass spectrometry. PRJ aided in data analysis. CJo and BS did experiments not included in the final manuscript and CJo constructed the Δ*ptkA* mutant. Manuscript was written by CJe and IM.

Joint author statement

If a thesis contains articles made in collaboration with other researchers, a joint author statement about the PhD-student's part of the article shall be made by each of the co-authors, cf. article 12, section 4 of the Ministerial Order No. 18 February 2008 about the PhD degree

Title of the article:	*Bacillus subtilis* two-component system sensory kinase DegS is regulated by serine phosphorylation in its input domain.		
Author(s):	Jers, C., Søndergaard, E.O., Kobir, A., Jensen, P.R. and Mijakovic, I.		
Journal:	Molecular microbiology (submitted)		
PhD-student:	Carsten Jers	CPR-no.:	180179-2835
Signature of the PhD-student:		Date:	10 May 2010
Co-author:	Elsebeth Oline Søndergaard	Signature:	
Co-author:	Ahasanul Kobir	Signature:	
Co-author:	Peter Ruhdal Jensen	Signature:	
Co-author:	Ivan Mijakovic	Signature:	Prof Dr Ivan MIJAKOVIĆ

Description of each author's contribution to the above-mentioned article:

Experiments were planned by IM and CJ. *in vitro* phosphorylation assays were done by EOS, CJ. AK and IM. AK purified Hanks type kinases. All strains and physiological assays were done by CJ. PRJ aided in data analysis. The manuscript was written by CJ and IM.

ISBN 978-87-9149-479-6